THE QUARTERLY JOURNAL
OF
MATHEMATICS
OXFORD SECOND SERIES

EDITED BY

T. W. CHAUNDY U. S. HASLAM-JONES
E. C. THOMPSON

VOLUME XII
1961

OXFORD
AT THE CLARENDON PRESS

Oxford University Press, Amen House, London E.C.4
GLASGOW NEW YORK TORONTO MELBOURNE WELLINGTON
BOMBAY CALCUTTA MADRAS KARACHI LAHORE DACCA
CAPE TOWN SALISBURY NAIROBI IBADAN ACCRA
KUALA LUMPUR HONG KONG

PRINTED IN GREAT BRITAIN AT THE UNIVERSITY PRESS, OXFORD
BY VIVIAN RIDLER, PRINTER TO THE UNIVERSITY

THE QUARTERLY JOURNAL OF MATHEMATICS

OXFORD SECOND SERIES

Volume 12 *No.* 45 *March* 1961

OXFORD

AT THE CLARENDON PRESS

***Price* 16*s.* net**

PRINTED IN GREAT BRITAIN BY VIVIAN RIDLER AT THE UNIVERSITY PRESS, OXFORD

THE QUARTERLY JOURNAL OF MATHEMATICS

OXFORD SECOND SERIES

Edited by T. W. CHAUNDY, U. S. HASLAM-JONES, E. C. THOMPSON

THE QUARTERLY JOURNAL OF MATHEMATICS (OXFORD SECOND SERIES) is published at 16*s.* net for a single number with an annual subscription (for four numbers) of 55*s.* post free.

Papers, of a length normally not exceeding 20 printed pages of the Journal, are invited on subjects of Pure and Applied Mathematics, and should be addressed 'The Editors, Quarterly Journal of Mathematics, Clarendon Press, Oxford'. *Authors are referred to 'The Printing of Mathematics' (Oxford University Press, 1954) for detailed advice on the preparation of mathematical papers for publication.* The Editors as a rule will not wish to accept material that they cannot see their way to publish within a twelvemonth.

While every care is taken of manuscripts submitted for publication, the Publisher and the Editors cannot hold themselves responsible for any loss or damage. Authors are advised to retain a copy of anything they may send for publication. Authors of papers printed in the Quarterly Journal will be entitled to 50 free offprints.

Correspondence on the *subject-matter* of the Quarterly Journal should be addressed, as above, to 'The Editors', at the Clarendon Press. All other correspondence should be addressed to the publishers:

OXFORD UNIVERSITY PRESS

AMEN HOUSE, LONDON, E.C. 4

MINIMUM MODULUS THEOREMS FOR FUNCTIONS REGULAR IN THE UNIT CIRCLE

By C. N. LINDEN (*Swansea*)

[Received 10 September 1959]

PART I

1. Introduction

LET $f(x)$ be non-negative for all large x and let

$$f_n(x) = \inf_{h>0} h^{-n}f(x+h). \tag{1.1}$$

Bounds to the function $f_n(x)$ have been established by Hayman and Stewart (1) who have proved the theorem:

THEOREM A. *Suppose that $f(x)$ and its derivatives up to order $n-1$ are non-negative and that $f^{(n-1)}(x)$ is a non-decreasing convex function for $x \geqslant X_0$. If $\epsilon > 0$, let E denote the set of x in (X_0, ∞) for which*

$$f_n(x) \leqslant (1+\epsilon)(e/n)^n f^{(n)}(x). \tag{1.2}$$

Then, if E_X denotes the subset of E contained in the interval (X_0, X), we have

$$\liminf_{X\to\infty} X^{-1}l(E_X) > C(\epsilon, n) > 0, \tag{1.3}$$

where $l(E_X)$ is the measure of the set E_X.

Hayman (2) has made application of Theorem A in examining the minimum modulus of integral functions. In this paper I shall be concerned not with integral functions which are regular everywhere in the finite part of the plane, but with functions regular only for $|z| < 1$. Thus in place of $f_n(x)$ I shall consider

$$g_n(1,r) = \inf_{r<R<1} (R-r)^{-n}g(R),$$

where $g(r)$ is a non-negative function defined for $r_0 \leqslant r < 1$.

In Part I, I shall deduce Theorem 1 for $g_n(1,r)$ from the corresponding Theorem A quoted above for $f_n(x)$ and then compare the results of the two theorems. In Part II, I shall use Theorem 1 in an examination of the minimum modulus of functions regular for $|z| < 1$, making applications of the methods used by Hayman (2) in his consideration of integral functions.

Throughout the paper C represents a positive constant not necessarily taking the same value at each occurrence and possibly depending on

one or more parameters which are generally indicated. Suffixes are used to distinguish distinct values C when this is desirable. We denote the measure of a set $\mathscr{E}$ of points on the real axis by $l(\mathscr{E})$.

2. Upper bounds to the function $g_n(1,r)$ will be given by the theorem:

THEOREM 1. *Suppose that $g(r)$ and its derivatives up to order $n-1$ are non-negative and that $g^{(n-1)}(r)$ is a non-decreasing convex function for $r_0 \leqslant r < 1$. Given $\epsilon > 0$, let $\mathscr{E}_{t,k}$ denote the set of values r in the interval $(t, (t+k)/(1+k))$ at which*

$$g_n(1,r) \leqslant (1+\epsilon)\left(\frac{e}{n}\right)^n \sum_{p=0}^{n} \frac{n!}{p!} g^{(p)}(r)(1-r)^{p-n}. \tag{2.1}$$

Then, if $1/k = 1/k(\epsilon, n)$ and $1-r_0'$ are sufficiently small positive numbers,

$$l(\mathscr{E}_{t,k}) > C(\epsilon, n)(1-t), \tag{2.2}$$

for $r_0' \leqslant t < 1$.

If a function $f(x)$ satisfies the hypothesis of Theorem A, then $f^{(n-1)}(x)$ has left and right derivatives which are equal except on an enumerable set, and each of these derivatives is a non-decreasing function of x. In the proof of Theorem A, $f^{(n)}(x)$ was used to denote the left derivative of $f^{(n-1)}(x)$, and I adopt the same notation here writing $g^{(n)}(r)$ for the left derivative of $g^{(n-1)}(r)$.

We deduce Theorem 1 as a corollary of Theorem A. Let us put

$$r = x/(1+x) \tag{2.3}$$

and

$$F(x) = g\left(\frac{x}{1+x}\right) \equiv g(r). \tag{2.4}$$

Then, if

$$g(r+h) = F(x+H),$$

we have

$$r+h = \frac{x+H}{1+x+H},$$

so that

$$h = \frac{H}{(1+x+H)(1+x)}.$$

It follows that

$$\begin{aligned} g_n(1,r) &= \inf_{0<h<1-r} h^{-n} g(r+h) \\ &= \inf_{H>0} H^{-n} F(x+H)(1+x+H)^n (1+x)^n. \end{aligned} \tag{2.5}$$

We shall assume for the moment that the function

$$f(x) = F(x)(1+x)^n \tag{2.6}$$

satisfies the hypothesis of Theorem A. Then

$$g_n(1,r) \leqslant (1+\epsilon)(e/n)^n f^{(n)}(x)(1+x)^n \tag{2.7}$$

for a set of values r which, by (2.3), correspond to the set of values E

of Theorem A. Now by (2.3), (2.4), and (2.6) we have that, if $m \leqslant n$, then

$$f^{(m)}(x) = \left\{(1-r)^2 \frac{d}{dr}\right\}^m \{g(r)(1-r)^{-n}\}$$

$$= \sum_{p=0}^{m} \frac{(n-p)!}{(n-m)!} {}^mC_p\, g^{(p)}(r)(1-r)^{p+m-n}, \tag{2.8}$$

the latter equality being established by induction on m. Hence writing $m = n$ in (2.8) we obtain (2.1) from (2.7).

Let E' denote the set of values r in the interval $(r_0, 1)$ for which (2.1) is valid, and let $E_{t,k}$ and $\mathscr{E}_{t,k}$ denote the subsets of E' in $(r_0, (t+k)/(1+k))$ and $(t, (t+k)/(1+k))$ respectively. Then, since E' is transformed by (2.3) into the set E of Theorem A, (1.3) implies that

$$\liminf_{t\to 1} \frac{1-t}{1+k} \int_{E_{t,k}} (1-r)^{-2}\, dr > C(\epsilon, n) = C > 0,$$

and, if $1-t$ is sufficiently small,

$$Ck(1-t)^{-1} < \int_{E_{t,k}} (1-r)^{-2}\, dr$$

$$\leqslant \int_{r_0}^{t} (1-r)^{-2}\, dr + \int_{\mathscr{E}_{t,k}} (1-r)^{-2}\, dr$$

$$< (1-t)^{-1} + (1+k)^2(1-t)^{-2} l(\mathscr{E}_{t,k}).$$

Hence $$l(\mathscr{E}_{t,k}) > (Ck-1)(1+k)^{-2}(1-t),$$

and (2.2) follows since, if $C > 0$, we can choose k depending on C so that $Ck-1 > 0$.

In order to complete the proof of Theorem 1 we must show that the function $f(x)$ defined by (2.6) satisfies the hypothesis of Theorem A. The equation (2.8) and the hypothesis of Theorem 1 show that $f(x)$ and all its derivatives up to order $n-1$ are non-negative. Further the function $f^{(n-1)}(x)$ can be differentiated on the right and left, and the two derivatives are equal except on an enumerable set. In particular the left derivative $f^{(n)}(x)$ is non-negative since $g(r), \ldots, g^{(n)}(r)$ are all non-negative, and this implies that $f^{(n-1)}(x)$ is a non-decreasing function of x. We prove that $f^{(n-1)}(x)$ is convex by proving that $f^{(n)}(x)$ is a non-decreasing function of x.

Let us write $$D = f^{(n)}(x_2) - f^{(n)}(x_1),$$

where $x_1 < x_2$, and suppose that r_1, r_2 correspond respectively to x_1, x_2

under the transformation (2.3). We shall see that $D \geqslant 0$. According to (2.8) with $m = n$ we have

$$D = \sum_{p=0}^{n} \frac{n!}{p!}\{g^{(p)}(r_2)(1-r_2)^p - g^{(p)}(r_1)(1-r_1)^p\}$$

$$\geqslant \sum_{p=0}^{n-1} \frac{n!}{p!}\Big\{g^{(p)}(r_2)(1-r_2)^p - g^{(p)}(r_1)\sum_{t=0}^{p} {}^pC_t(1-r_2)^t(r_2-r_1)^{p-t}\Big\} -$$
$$- g^{(n)}(r_1)\sum_{t=0}^{n-1} {}^nC_t(1-r_2)^t(r_2-r_1)^{n-t}$$

since $g^{(n)}(r)$ is a non-decreasing function of r. Collecting together coefficients of $(1-r_2)^p$ we rewrite this inequality as

$$D \geqslant \sum_{p=0}^{n-1} \frac{n!}{p!}(1-r_2)^p\left\{g^{(p)}(r_2) - \sum_{s=0}^{n-p} \frac{(r_2-r_1)^s}{s!} g^{(p+s)}(r_1)\right\}. \tag{2.9}$$

By Taylor's theorem we see that, for a given p $(0 \leqslant p \leqslant n-1)$, there exists M such that $g^{(n)}(r_1) \leqslant M \leqslant g^{(n)}(r_2)$ and

$$g^{(p)}(r_2) = \sum_{s=0}^{n-p-1} \frac{(r_2-r_1)^s}{s!} g^{(p+s)}(r_1) + \frac{M}{(n-p)!}(r_2-r_1)^{n-p}$$

$$\geqslant \sum_{s=0}^{n-p} \frac{(r_2-r_1)^s}{s!} g^{(p+s)}(r_1).$$

It follows that each of the n terms in the summation on the right-hand side of (2.9) is non-negative, and so $D \geqslant 0$ as required.

3. In Theorem A we see that the right-hand side of (1.2) contains the derivative $f^{(n)}(x)$ but no other term depending on $f(x)$, whereas the right-hand side of (2.1) in Theorem 1 consists of the sum of multiples of $g^{(n)}(r)$, $g^{(n-1)}(r)(1-r)^{-1}, \ldots, g(r)(1-r)^{-n}$. Consequently we might inquire whether (2.1) may be replaced by an inequality of the form

$$g_n(1, r) < Cg^{(n)}(r).$$

Theorem 2 shows that this is not possible in general and that any upper bound to $g_n(1, r)$ must contain both of the terms $g^{(n)}(r)$ and $g(r)(1-r)^{-n}$.

THEOREM 2. (i) *If $g(r) = -\log(1-r)$, then as r increases to 1 we have*

$$g_n(1, r) \sim (1-r)^{-n}g(r) \tag{3.1}$$

and
$$(1-r)^{p-n}g^{(p)}(r) = o\{(1-r)^{-n}g(r)\}, \tag{3.2}$$
for $1 \leqslant p \leqslant n$.

(ii) *If* $g(r) = \exp\{(1-r)^{-1}\}$, *then as r increases to* 1 *we have*

$$g_n(1,r) \sim \left(\frac{e}{n}\right)^n g^{(n)}(r) \tag{3.3}$$

and
$$(1-r)^{p-n}g^{(p)}(r) = o\{g^{(n)}(r)\}, \tag{3.4}$$
for $0 \leqslant p \leqslant n-1$.

Proof. (i) If $p \geqslant 1$, then

$$g^{(p)}(r) = (p-1)!(1-r)^{-p}.$$

Hence
$$(1-r)^{p-n}g^{(p)}(r) = (p-1)!(1-r)^{-n},$$
and we have (3.2) for $1 \leqslant p \leqslant n$.

If $0 < h < 1-r$, then

$$-h^{-n}\log(1-r-h) \geqslant -(1-r)^{-n}\log(1-r),$$

and therefore
$$g_n(1,r) \geqslant (1-r)^{-n}g(r). \tag{3.5}$$
On the other hand suppose that $0 < \theta < 1$ and set

$$h = (1-\theta)(1-r).$$

Then, if $\epsilon > 0$,

$$\begin{aligned} g_n(1,r) &\leqslant -(1-\theta)^{-n}(1-r)^{-n}\{\log\theta+\log(1-r)\} \\ &\leqslant (1+\epsilon)(1-r)^{-n}g(r) \end{aligned} \tag{3.6}$$

provided that $(1-\theta)^{-n} < 1+\epsilon$ and $r > r_0(\theta)$. Since ϵ can be chosen arbitrarily small, (3.1) follows from (3.5) and (3.6).

(ii) We estimate $g_n(1,r)$ by obtaining the minimum value of the function

$$G(h) = h^{-n}g(r+h)$$

when $0 < h < 1-r$. At the stationary values of $G(h)$ we have

$$\frac{n}{h} = \frac{g'(r+h)}{g(r+h)} = (1-r-h)^{-2},$$

and it is seen that, as $r \to 1-0$, $G(h)$ has a minimum value for

$$h \sim n(1-r)^2.$$

Thus

$$g_n(1,r) = \inf_{0<h<1-r} h^{-n}g(r+h) \sim (e/n)^n(1-r)^{-2n}\exp\{(1-r)^{-1}\},$$

and we have (3.3).

The relation (3.4) is easily verified since

$$(1-r)^{p-n}g^{(p)}(r) \sim (1-r)^{-p-n}g(r) = o\{g^{(n)}(r)\},$$

as $r \to 1-0$ when $0 \leqslant p \leqslant n-1$.

Part II

4. The association which exists between the minimum modulus

$$\mu(r,f) = \min_{|z|=r} |f(z)|$$

and the maximum modulus

$$M(r,f) = \max_{|z|=r} |f(z)|$$

of a function $f(z)$ regular for $|z| < R$ has been considered by several authors. In particular, when R is infinite, $f(z)$ represents an integral function, and Hayman (**2**) has shown that for such functions we have

$$\limsup_{r\to\infty} \frac{\log \mu(r,f)}{\log M(r,f) \log\log\log M(r,f)} > -2{\cdot}19. \tag{4.1}$$

If the order
$$\rho = \limsup_{r\to\infty} \frac{\log\log M(r,f)}{\log r}$$

is finite, (4.1) can be sharpened to

$$\limsup_{r\to\infty} \frac{\log \mu(r,f)}{\log M(r,f)} > -C(\rho) > -\infty, \tag{4.2}$$

and both results are best-possible.

We shall be concerned here with the case in which R is finite and we shall consider functions $f(z)$ regular for $|z| < 1$. If we define

$$\beta = \limsup_{r\to 1-0} \frac{\log\log M(r,f)}{-\log(1-r)}$$

as the order of $f(z)$ in $|z| < 1$, then it has been shown (**3**) that, if $1 < \beta < \infty$, then

$$\limsup_{r\to 1-0} \frac{\log \mu(r,f)}{\log M(r,f) \log\log M(r,f)} > -C(\beta) > -\infty, \tag{4.3}$$

and that this result cannot be improved in general. We shall now prove that, whenever $f(z)$ is of infinite order in $|z| < 1$, we can replace $C(\beta)$ by an arbitrary positive constant, and (4.3) still remains true. Further, if the rate of growth of $M(r,f)$ is sufficiently rapid, the relation between $\mu(r,f)$ and $M(r,f)$ can be sharpened and an inequality of the form (4.1) is again valid. We prove the theorems:

Theorem 3. *If $f(z)$ is regular and of infinite order in $|z| < 1$, then*

$$\limsup_{r\to 1-0} \frac{\log \mu(r,f)}{\log M(r,f) \log\log M(r,f)} = 0. \tag{4.4}$$

THEOREM 4. *Suppose that $f(z)$ is regular in $|z| < 1$ and that*

$$\limsup_{r\to 1-0} \frac{\log\log\log M(r,f)}{-\log(1-r)} \geqslant \sigma, \tag{4.5}$$

where $0 < \sigma < \infty$. Then we have

$$\limsup_{r\to 1-0} \frac{\log \mu(r,f)}{\log M(r,f)\log\log\log M(r,f)} \geqslant -C\left(1+\frac{1}{\sigma}\right). \tag{4.6}$$

It should be noted that, although σ is finite in (4.5), the limit on the left-hand side may be infinite. Further the inequality (4.6) cannot be sharpened in general since Hayman (**2**) has given examples of functions regular in $|z| < 1$ for which the maximum modulus increases arbitrarily rapidly as $r \to 1-0$ and for which

$$\log \mu(r,f) < -C \log M(r,f)\log\log\log M(r,f),$$

where C is an absolute positive constant.

5. In verifying the inequalities (4.1) and (4.2) Hayman (**2**) has written the function $u(z) = \log|f(z)|$ as

$$u(z) = u(z,h) - N(z,h), \tag{5.1}$$

where
$$N(z,h) = \int_0^h t^{-1} n(z,t)\,dt,$$

and $n(z,t)$ is the number of zeros of the function $f(\xi)$ in the circle $|z-\xi| \leqslant t$. Hayman has also defined

$$B_2(r,f) = \int_0^r (r-t)B(t,f)\,dt,$$

where
$$B(t,f) = \log M(t,f).$$

We use this notation in quoting the following results from his paper (**2**):

THEOREM B. *Suppose that $f(z)$ is regular for $|z| \leqslant R < \infty$, $f(0) = 1$ and that $0 < r < R$. Then there exist positive constants η, C_1, and C_2 such that $\eta < 1$ and, for $|z| = r$, $h = \eta(R-r)$ we have*

$$u(z, \tfrac{1}{2}h) \geqslant -C_1(R-r)^{-2}B_2(R,f), \tag{5.2}$$

$$n(z,h) \leqslant C_2(R-r)^{-2}B_2(R,f). \tag{5.3}$$

LEMMA A. *Suppose that $r > 0$, $h > 0$ and that for $|z| = r$ we have $n(z,h) \leqslant n_0$. Then there exists a set $\mathscr{E}$ of ρ for which $r < \rho < r+\frac{1}{2}h$ and having measure at least $\frac{1}{4}h$ such that, for ρ in $\mathscr{E}$ and $|z| = \rho$,*

$$N(z, \tfrac{1}{2}h) \leqslant n_0 \log \frac{C(r+h)}{h}.$$

According to the above definitions we have

$$B_2^{(1)}(r,f) = \int_0^r B(t,f)\,dt, \qquad B_2^{(2)}(r,f) = B(r,f).$$

It is readily seen that, if $f(0) = 1$, then $B_2^{(2)}(r,f)$, $B_2^{(1)}(r,f)$, and $B_2(r,f)$ are non-negative, non-decreasing functions of r and that

$$B_2^{(2)}(r,f) \geqslant B_2^{(1)}(r,f) \geqslant B_2(r,f),$$

for $0 \leqslant r < 1$. For convenience we shall generally write

$$B_2(r,f) = B_2(r), \qquad B_2^{(1)}(r,f) = B_2^{(1)}(r), \quad \text{etc.}$$

6. Hayman (2) has combined Theorem A with Theorem B and Lemma A to obtain lower bounds to $|f(z)|$ on a certain set of circles when $f(z)$ is an integral function. We combine the corresponding Theorem 1 with Theorem B and Lemma A to prove the following lemma for functions regular in $|z| < 1$.

LEMMA 1. *Suppose that $f(z)$ is regular and non-constant for $|z| < 1$ and that $f(0) = 1$. Then there is a positive constant k such that in each interval $(t, (t+k)/(1+k))$, where $t_0 \leqslant t < 1$, there is a set of r of measure at least $C_3(1-t)$ such that*

$$\log|f(z)| > -CP(r)\log\frac{P(r)}{B_2(r)}, \tag{6.1}$$

on $|z| = r$, where

$$P(r) = B_2^{(2)}(r)+2(1-r)^{-1}B_2^{(1)}(r)+2(1-r)^{-2}B_2(r).$$

We note that, when $n = 2$, the function $g(r) = B_2(r)$ satisfies the hypothesis of Theorem 1 since $B_2(r)$, $B_2^{(1)}(r)$, and $B_2^{(2)}(r)$ are all non-negative, and $B_2^{(2)}(r) = \log M(r,f)$ is an increasing function of r. It follows that, if we take $\epsilon = 1$ and $n = 2$, the constant k of Theorem 1 is an absolute constant, and we have that, if $t_0 \leqslant t < 1$, then in each interval $(t, (t+k)/(1+k))$ there is a set $\mathscr{E}_{t,k}$ of r having measure at least $C(1-t)$ such that, for the corresponding values R,

$$(R-r)^{-2}B_2(R) \leqslant \tfrac{1}{2}e^2P(r). \tag{6.2}$$

Since $f(z)$ is regular for $|z| \leqslant R < 1$, the application of Theorem B gives

$$n(z,h) \leqslant C_2(R-r)^{-2}B_2(R) \leqslant \tfrac{1}{2}e^2C_2\,P(r),$$

on $|z| = r$, where $h = \eta(R-r)$ and $0 < \eta < 1$. Lemma A now yields

$$N(z,\tfrac{1}{2}h) \leqslant \tfrac{1}{2}e^2C_2\,P(r)\log\frac{C}{h} < -CP(r)\log(R-r) \tag{6.3}$$

on $|z| = \rho$ for a set of ρ in $(r, r+\frac{1}{2}h)$ of measure at least $\frac{1}{4}h$.

We consider a value ρ for which (6.3) holds on $|z| = \rho$. Then

$$r < \rho < r+\tfrac{1}{2}h$$

and $$R-\rho > R-r-\tfrac{1}{2}h = (R-r)(1-\tfrac{1}{2}\eta) > \tfrac{1}{2}(R-r).$$

A second application of Theorem B to the function $f(z)$ shows that on $|z| = \rho$ we have

$$\begin{aligned} u(z, \tfrac{1}{2}h') &\geqslant -C_1(R-\rho)^{-2}B_2(R) \\ &> -4C_1(R-r)^2B_2(R) \\ &> -CP(r) \end{aligned}$$

for $h' = \eta(R-\rho)$. Since $u(z,h)$ is a non-decreasing function of h, we deduce

$$u(z, \tfrac{1}{2}h) \geqslant u(z, \tfrac{1}{2}h') > -CP(r),$$

and hence $$u(z) = u(z, \tfrac{1}{2}h)-N(z, \tfrac{1}{2}h) > CP(r)\log(R-r) \tag{6.4}$$

on $|z| = \rho$, by (6.3).

The inequality (6.2) gives

$$(R-r)^{-2} < \tfrac{1}{2}e^2\frac{P(r)}{B_2(R)} < \tfrac{1}{2}e^2\frac{P(\rho)}{B_2(\rho)} \tag{6.5}$$

since both the functions $P(r)$ and $B_2(r)$ increase with r. Hence combining (6.4) with (6.5) we have

$$u(z) > -CP(r)\log\frac{P(\rho)}{B_2(\rho)},$$

so that $$u(z) > -CP(\rho)\log\frac{P(\rho)}{B_2(\rho)} \tag{6.6}$$

on $|z| = \rho$.

We complete the proof of Lemma 1 by showing that the set of ρ in $(t, (t+k)/(1+k))$ for which (6.6) holds has measure at least $C(1-t)$. The proof is that of Hayman (**2**). The first of the inequalities (6.5) shows that $R-r$ has a positive lower bound, say δ, when

$$t \leqslant r \leqslant (t+k)/(1+k).$$

We choose a sequence of intervals (r_n, R_n) such that

(i) $R_0 = t < r_1 < R_1 < ... < r_n < R_n < r_{n+1} < ...$;

(ii) r_n lies in the set $\mathscr{E}_{t,k}$ of r for which (6.2) holds and R_n is the corresponding value R for $n > 0$;

(iii) the measure of the part of $\mathscr{E}_{t,k}$ in (R_n, r_{n+1}) is less than δ for $n \geqslant 0$.

Let N be the largest integer such that $r_N < (t+k)/(1+k)$. Then the measure of that part of $\mathscr{E}_{t,k}$ in the intervals (R_n, r_{n+1}) belonging to $(t, (t+k)/(1+k))$ is at most $(N+1)\delta$, and the total measure of $\mathscr{E}_{t,k}$ is at most

$$(N+1)\delta + \sum_{n=1}^{N}(R_n-r_n) = (N+1)\delta+l,$$

say. Now, by hypothesis, $R_n - r_n \geqslant \delta$, so that $l \geqslant N\delta$ and the measure of $\mathscr{E}_{t,k}$ is at most $(2+N^{-1})l$. However, $\mathscr{E}_{t,k}$ has measure at least $C(1-t)$ by Theorem 1, and hence $l > C(1-t)$.

The inequality (6.6) holds for a set of ρ of measure at least $\frac{1}{4}(R_n - r_n)$ in each interval (r_n, R_n). Hence the values ρ in the interval

$$(t, (t+k)/(1+k))$$

for which (6.6) holds on $|z| = \rho$ have measure at least $\frac{1}{4}l > C(1-t)$. This is equivalent to the statement of Lemma 1.

7. We shall deduce Theorem 3 from Lemma 1 by obtaining upper bounds to the functions $P(r)$ and $P(r)/B_2(r)$ which occur in the right-hand side of (6.1). We consider first the function $P(r)$ obtaining upper bounds when $B_2(r)$ is fairly large, proving the lemma:

LEMMA 2. *If $f(z)$ is regular and of infinite order in $|z| < 1$, then, for each $K > 1$, there is a sequence of values t increasing to 1 such that*

$$B_2(t) > (1-t)^{-K} \tag{7.1}$$

and

$$P(r) < C(k) B_2^{(2)}(r) \tag{7.2}$$

for $t \leqslant r \leqslant (t+k)/(1+k)$.

The proof of Lemma 2 is based on two lemmas. The first of these is

LEMMA 3. *If $f(z)$ is regular and of infinite order in $|z| < 1$, then, for each $K > 0$, there is a sequence of values τ increasing to 1 such that*

$$B_2(\tau) > (1-\tau)^{-K} \tag{7.3}$$

and

$$\frac{B_2^{(1)}(\tau)}{B_2(\tau)} \geqslant \frac{K}{1-\tau}. \tag{7.4}$$

We begin by showing that

$$\limsup_{r \to 1-0} \frac{\log B_2(r)}{-\log(1-r)} = \infty. \tag{7.5}$$

If (7.5) is false, there are finite constants A and C such that

$$B_2(r) < A(1-r)^{-C}$$

for $0 \leqslant r < 1$. However, by definition,

$$B_2(r) \geqslant \int_{2r-1}^{r} (r-t) B(t)\, dt \geqslant \tfrac{1}{2}(1-r)^2 B(2r-1)$$

since $B_2(t)$ is an increasing function of t. Hence

$$\log M(2r-1) = B(2r-1) \leqslant 2A(1-r)^{-C-2},$$

and $f(z)$ is of finite order, contrary to the hypothesis of Lemma 3. Therefore (7.5) must be true.

By (7.5) we have that, for any constant K, there is a sequence of values r increasing to 1 such that

$$B_2(r) > (1-r)^{-2K}. \tag{7.6}$$

Consequently, for some value t in the interval $(0, r)$,

$$\frac{B_2^{(1)}(t)}{B_2(t)} \geqslant \frac{2K}{1-t}.$$

For, if not,

$$\log B_2(r) = \int_0^r \frac{B_2^{(1)}(t)}{B_2(t)}\,dt < -2K\log(1-r),$$

which contradicts (7.6).

If

$$\frac{B_2^{(1)}(r)}{B_2(r)} > \frac{K}{1-r}, \tag{7.7}$$

then (7.4) is satisfied, and (7.6) implies (7.3) with $\tau = r$. If (7.7) is false, we define an interval (r', r) such that

$$\frac{B_2^{(1)}(r')}{B_2(r')} = \frac{K}{1-r'}$$

and

$$\frac{B_2^{(1)}(t)}{B_2(t)} < \frac{K}{1-t},$$

for $r' < t < r$. Then, by integration we have

$$\log\frac{B_2(r)}{B_2(r')} < K\log\frac{1-r'}{1-r} < K\log\frac{1}{1-r}.$$

It follows from (7.6) that

$$\log B_2(r') > \log B_2(r) + K\log(1-r) > -K\log(1-r).$$

Thus, if r increases to 1, the corresponding value r' tends to 1, and we have further

$$\log B_2(r') > -K\log(1-r').$$

We have proved that, for the given value r, the inequalities (7.3) and (7.4) are both valid either with $\tau = r$ or with $\tau = r'$. The sequence of values r increases to 1, and the sequence of values r', when it exists, tends to 1. Hence we can select a sequence of values τ increasing to 1 for which (7.3) and (7.4) are true simultaneously.

8. In this section we prove the lemma:

LEMMA 4. *If $f(z)$ is regular and of infinite order in $|z| < 1$, then, for each $K > 0$, there is a sequence of values T increasing to 1 for which*

$$B_2^{(1)}(T) > (1-T)^{-K}, \tag{8.1}$$

$$(1-T)^2 B_2^{(2)}(T) \geqslant K\{(1-T)B_2^{(1)}(T) + B_2(T)\}. \tag{8.2}$$

Consider a value τ for which (7.3) and (7.4) hold. Then there is a value r in the interval $(0,\tau)$ such that

$$B_2^{(2)}(r) \geqslant \tfrac{1}{2}K\{(1-r)^{-1}B_2^{(1)}(r)+(1-r)^{-2}B_2(r)\}. \tag{8.3}$$

For, if not, integration over the range $(0,\tau)$ gives

$$B_2^{(1)}(\tau) < \tfrac{1}{2}K(1-\tau)^{-1}B_2(\tau),$$

which contradicts (7.4).

If (8.3) holds with $r = \tau$, then (8.2) is true with $T = \tau$ and K replaced by $\frac{1}{2}K$; so is (8.1) since (7.3) implies

$$B_2^{(1)}(\tau) > B_2(\tau) > (1-\tau)^{-K} > (1-\tau)^{-\frac{1}{2}K}.$$

Thus, we have the required inequalities with a change of notation.

If (8.3) is false when $r = \tau$, we define an interval (T,τ) such that (8.3) is false for $T < r < \tau$ and

$$B_2^{(2)}(T) = \tfrac{1}{2}K\{(1-T)^{-1}B_2^{(1)}(T)+(1-T)^{-2}B_2(T)\}. \tag{8.4}$$

By integration we have

$$B_2^{(1)}(\tau)-B_2^{(1)}(T) < \tfrac{1}{2}K\{(1-\tau)^{-1}B_2(\tau)-(1-T)^{-1}B_2(T)\},$$

and therefore

$$\begin{aligned} B_2^{(1)}(T) &> B_2^{(1)}(\tau)-\tfrac{1}{2}K(1-\tau)^{-1}B_2(\tau) \\ &\geqslant \tfrac{1}{2}B_2^{(1)}(\tau) \\ &> \tfrac{1}{2}B_2(\tau) \\ &> \tfrac{1}{2}(1-\tau)^{-K}, \end{aligned}$$

by (7.3) and (7.4). It follows that, if τ increases to 1, then T must tend to 1 and, choosing $1-\tau$ sufficiently small, we have

$$B_2^{(1)}(T) > (1-\tau)^{-\frac{1}{2}K} > (1-T)^{-\frac{1}{2}K}. \tag{8.5}$$

The inequalities (8.4) and (8.5) give (8.1) and (8.2) with K again replaced by $\frac{1}{2}K$. This substitution can be made without any loss of generality, and we have (8.1) and (8.2) for a sequence of T increasing to 1.

9. In the proof of Lemma 2 we consider the function

$$Q(r) = \frac{B_2^{(1)}(r)+(1-r)^{-1}B_2(r)}{B_2^{(2)}(r)},$$

supposing that $r > T$, where T is a value for which (8.1) and (8.2) are true. Then, since $B_2^{(2)}(r)$ and $B_2^{(1)}(r)$ are both increasing functions of r,

the Taylor series gives

$$\begin{aligned}
&B_2^{(1)}(r)+(1-r)^{-1}B_2(r)\\
&\qquad < B_2^{(1)}(T)+(r-T)B_2^{(2)}(r)+(1-r)^{-1}\{B_2(T)+(r-T)B_2^{(1)}(r)\}\\
&\qquad = B_2^{(1)}(T)+(1-T)^{-1}B_2(T)+\\
&\qquad\quad +(r-T)\{B_2^{(2)}(r)+(1-r)^{-1}B_2^{(1)}(r)+(1-r)^{-1}(1-T)^{-1}B_2(T)\}\\
&\qquad < B_2^{(2)}(T)Q(T)+(r-T)B_2^{(2)}(r)\{1+(1-r)^{-1}Q(r)\}.
\end{aligned}$$

Hence $$Q(r) < Q(T)+(r-T)\{1+(1-r)^{-1}Q(r)\}.$$

Now (8.2) shows that

$$Q(T) \leqslant \frac{1-T}{K} < 1-T$$

since $K > 1$, and, if we suppose that

$$T < r \leqslant \tfrac{1}{4}(3T+1),$$

then $$\frac{r-T}{1-r} \leqslant \tfrac{1}{3},$$

so that $$Q(r) < \tfrac{3}{2}\{1-T+r-T\} < 3(1-T) < 4(1-r).$$

We define a sequence of values $\{T_i\}$ such that

$$T_0 = T \quad\text{and}\quad T_{i+1} = \tfrac{1}{4}(3T_i+1) \quad\text{for } i = 0,1,2,\ldots.$$

Repeating the above argument we find that, if r lies in the interval (T_i, T_{i+1}), then

$$Q(r) < 4^i(1-r). \tag{9.1}$$

By definition we have

$$T_{i+1} = 1-(\tfrac{3}{4})^i(1-T_1),$$

and given k we choose i to be the smallest integer such that

$$T_{i+1} \geqslant (T_1+k)/(1+k),$$

that is $$(\tfrac{4}{3})^i \geqslant 1+k.$$

It follows that, if $t = T_1$, then (9.1) holds for $t \leqslant r \leqslant (t+k)/(1+k)$. Thus we have

$$Q(r) \leqslant C(k)(1-r),$$

and (7.2) follows by definition of $Q(r)$ and $P(r)$.

We have selected T to satisfy (8.1). Hence

$$\begin{aligned}
B_2(t) &> B_2(T)+B_2^{(1)}(T)(t-T)\\
&> \tfrac{1}{4}(1-T)^{-K+1}\\
&\geqslant \tfrac{1}{4}(\tfrac{3}{4})^{K-1}(1-t)^{-K+1}.
\end{aligned}$$

If $1-t$ is sufficiently small, this gives

$$B_2(t) > (1-t)^{-\frac{1}{2}(K-1)},$$

and we have (7.1) with a change of notation. Since T is one of a sequence of values increasing to 1, so is t and Lemma 2 is proved.

10. According to Lemma 2 we have certain intervals in which

$$P(r) < C(k)B_2^{(2)}(r).$$

In these intervals upper bounds to the function $P(r)/B_2(r)$ can readily be deduced from upper bounds to $B_2^{(2)}(r)/B_2(r)$. We find upper bounds to this latter function which are valid except possibly on an exceptional set. This exceptional set has small measure if the constant K of Lemma 2 is chosen sufficiently large. In fact we have the lemma:

LEMMA 5. *Suppose that $f(z)$ is regular in $|z| < 1$, that $B_2(r) \to \infty$ as $r \to 1-0$, and that $0 < \nu < 1$. Then the set of values r for which*

$$\{B_2^{(2)}(r)\}^{1-\nu} < B_2(r) \tag{10.1}$$

is false in $(t, 1)$ has measure at most $6\nu^{-1}B_2(t)^{-\frac{1}{3}\nu}$.

Suppose that $0 < \lambda < 1$ and let $\mathscr{E}_t$ denote the set of values in the interval $(t, 1)$ at which

$$B_2^{(1)}(r) > \{B_2(r)\}^{1+\lambda}.$$

Then $$l(\mathscr{E}_t) = \int_{\mathscr{E}_t} dr < \int_t^1 \{B_2(r)\}^{-1-\lambda}B_2^{(1)}(r)\,dr = \frac{1}{\lambda}\{B_2(t)\}^{-\lambda}.$$

Similarly the set of values in $(t, 1)$ at which

$$B_2^{(2)}(r) > \{B_2^{(1)}(r)\}^{1+\lambda}$$

has measure less than

$$\frac{1}{\lambda}\{B_2^{(1)}(t)\}^{-\lambda} \leqslant \frac{1}{\lambda}\{B_2(t)\}^{-\lambda}.$$

Since $0 < \lambda < 1$, we have $(1+\lambda)^2 < 1+3\lambda$ and thus

$$B_2^{(2)}(r) < \{B_2^{(1)}(r)\}^{1+\lambda} < \{B_2(r)\}^{1+3\lambda}, \tag{10.2}$$

except in a set of measure less than

$$2\lambda^{-1}\{B_2(t)\}^{-\lambda}.$$

If we choose $\lambda = \frac{1}{3}\nu$, then (10.2) yields (10.1). The measure of the set on which (10.1) is false is less than

$$2\lambda^{-1}\{B_2(t)\}^{-\lambda} = 6\nu^{-1}\{B_2(t)\}^{-\frac{1}{3}\nu}.$$

11. In proving Theorem 3 we suppose that $f(0) = 1$ since a proof of the theorem with this condition implies that the theorem is true generally. We consider one of the intervals $(t, (t+k)/(1+k))$ for which (7.1) and (7.2) hold, choosing K to satisfy the inequality

$$(1-t)^{\frac{1}{3}(K\nu-3)} < \tfrac{1}{12}\nu C_3 \tag{11.1}$$

for $t_0 \leqslant t < 1$. Here we take C_3, t_0, and ν to be the values assumed in Lemmas 1, 2, and 5 respectively.

By Lemma 1, for a set of values r of measure $C_3(1-t)$ in the interval $(t, (t+k)/(1+k))$ we have

$$\log|f(re^{i\theta})| > -CP(r)\log\frac{P(r)}{B_2(r)}, \tag{11.2}$$

for $0 \leqslant \theta < 2\pi$. The inequality (7.2) states that

$$P(r) < C(k)B_2^{(2)}(r), \tag{11.3}$$

and, since k is an absolute constant, we have

$$\frac{P(r)}{B_2(r)} < \frac{CB_2^{(2)}(r)}{B_2(r)} < C\{B_2^{(2)}(r)\}^\nu, \tag{11.4}$$

by application of Lemma 5, except for a set of r in $(t, 1)$ of measure less than

$$6\nu^{-1}\{B_2(t)\}^{-\frac{1}{2}\nu} < 6\nu^{-1}(1-t)^{\frac{1}{2}K\nu} < \tfrac{1}{2}C_3(1-t).$$

It follows that (11.2), (11.3), and (11.4) all hold in a set of measure at least $\frac{1}{2}C_3(1-t)$ in the interval $(t, (t+k)/(1+k))$. The combination of these three inequalities gives

$$\log|f(re^{i\theta})| > -\nu CB_2^{(2)}(r)\log B_2^{(2)}(r) \quad (0 \leqslant \theta < 2\pi),$$

which is

$$\log\mu(r,f) > -\nu C\log M(r,f)\log\log M(r,f). \tag{11.5}$$

The value t above is any one of a sequence which increases to 1. Hence (11.5) holds for a sequence of values r increasing to 1 and we have

$$\limsup_{r\to 1-0}\frac{\log\mu(r,f)}{\log M(r,f)\log\log M(r,f)} \geqslant -\nu C.$$

Since ν is any value such that $0 < \nu < 1$, we have proved that

$$\limsup_{r\to 1-0}\frac{\log\mu(r,f)}{\log M(r,f)\log\log M(r,f)} \geqslant 0. \tag{11.6}$$

It is clear that $\mu(r,f) \leqslant M(r,f)$ and therefore

$$\frac{\log\mu(r,f)}{\log M(r,f)\log\log M(r,f)} \leqslant \frac{1}{\log\log M(r,f)}.$$

Thus we have

$$\limsup_{r\to 1-0}\frac{\log\mu(r,f)}{\log M(r,f)\log\log M(r,f)} \leqslant 0. \tag{11.7}$$

The relation (4.4) of Theorem 3 follows from (11.6) and (11.7).

12. We shall not consider the proof of Theorem 4 in any detail since it can be proved in much the same manner as Theorem 3. However, we note the following two lemmas which supersede Lemmas 2 and 5 and which have similar proofs.

LEMMA 6. *Suppose that $f(z)$ is regular in $|z| < 1, f(0) = 1$, and that*

$$\limsup_{r \to 1-0} \frac{\log\log B_2(r)}{-\log(1-r)} \geqslant \sigma,$$

where $0 < \sigma < \infty$. Then, if $0 < \delta < \sigma$, there is a sequence of values t increasing to 1 such that

$$\log B_2(t) > (1-t)^{-\sigma+\delta},$$

and

$$P(r) < C(k) B_2^{(2)}(r), \tag{12.1}$$

for $t \leqslant r \leqslant (t+k)/(1+k)$.

LEMMA 7. *Suppose that $f(z)$ is regular for $|z| < 1$, $B_2(r) \to \infty$ as $r \to 1-0$, and $1 < \nu < \infty$, $\epsilon > 0$. Then, if $1-t$ is sufficiently small, we have*

$$B_2^{(2)}(r) < (1+\epsilon) B_2(r) (\log B_2(r))^{2\nu},$$

in $(t, 1)$ except for a set of values r of measure at most

$$\frac{2}{\nu - 1} \{\log B_2(t)\}^{1-\nu}.$$

Lemmas 6 and 7 show that, if $f(z)$ satisfies the hypothesis of Theorem 4 and $\sigma(\nu-1) > 1$, there is a sequence of intervals $(t_n, 1)$ such that $t_n \to 1-0$ as $n \to \infty$ and

$$\log \frac{P(r)}{B_2(r)} < \nu C \log\log B_2(r) \tag{12.2}$$

in $(t_n, 1)$ except for a set of values r of measure small in comparison with $1-t_n$. The inequality (4.6) can be deduced from (12.1) and (12.2) by the same procedure that was used to obtain (11.6) from (11.3) and (11.4).

I wish to record my thanks to the referee who made some helpful suggestions.

REFERENCES

1. W. K. Hayman and F. M. Stewart, *Proc. Cambridge Phil. Soc.* 50 (1954) 250–60.
2. W. K. Hayman, *Proc. London Math. Soc.* (3) 2 (1952) 469–512.
3. C. N. Linden, *Quart. J. of Math.* (Oxford) (2) 7 (1956) 196–216.

THE DETERMINATION OF PHASE SHIFT

By J. B. McLEOD (*Oxford*)

[Received 15 February 1960]

1. In the case of a wave function with spherical symmetry, the wave equation can be separated by using spherical polar coordinates, and the equation for the radial component becomes

$$\frac{d^2\psi}{dr^2}+\left\{\lambda-q(r)-\frac{l(l+1)}{r^2}\right\}\psi = 0, \tag{1.1}$$

where λ is a constant parameter, proportional to the energy of the particle under consideration, $q(r)$ is proportional to the potential energy, and l is a positive integer or zero.

If $q(r) \to 0$ as $r \to \infty$, and if $\lambda > 0$, then it is reasonable to suppose that, as $r \to \infty$, the solutions of (1.1) will be asymptotic to the solutions of

$$\frac{d^2\psi}{dr^2}+\lambda\psi = 0,$$

i.e. to $A\cos r\surd\lambda + B\sin r\surd\lambda$, where A, B are arbitrary constants. In fact, in order to achieve a rigorous mathematical proof of this result, it is necessary to impose further conditions on $q(r)$ than merely that it tend to zero.

Let us suppose that these further conditions, whatever they may be, have been imposed. Then the particular solution of (1.1) which is of physical interest is that for which $\psi = 0$ at $r = 0$, and, as $r \to \infty$, this solution must, apart from an unimportant constant factor, be asymptotic to

$$\cos\{r\surd\lambda - \tfrac{1}{2}\pi(l+1) - \eta_l\}, \tag{1.2}$$

for some constant η_l. (The reason for the introduction of the term $\frac{1}{2}\pi(l+1)$ will be evident later.) This constant η_l is called the *phase shift*, and its determination in any particular problem is of some physical interest. It is to be noted that the definition given of phase shift leaves it undetermined within a multiple of π; I shall choose the particular multiple of π which leads to the simplest expression for η_l, and this will be done later.

Estimates for η_l are to be found in the literature, e.g. [(1) equation

(9.3.109)]. There we find the formula (in our notation)

$$\eta_l = \int\limits_{(l+\frac{1}{2})/\surd\lambda}^{\infty} \left\{\lambda - \frac{(l+\frac{1}{2})^2}{r^2}\right\}^{\frac{1}{2}} dr - \int\limits_{r_0}^{\infty} \left\{\lambda - q(r) - \frac{(l+\frac{1}{2})^2}{r^2}\right\}^{\frac{1}{2}} dr, \tag{1.3}$$

where r_0 is the zero (assumed to be the only one) of the integrand in the corresponding integral. Since the two integrals do not separately converge, it is understood that the integrands are first to be subtracted and the difference then integrated over the infinite range.

The proof of this formula does not pretend to be rigorous, and no estimate is given of the errors involved. The aim of this paper is to provide a rigorous mathematical proof of (1.3), together with an exact mathematical expression for the error term. Of course, in any particular computation of η_l, it would probably be sufficient to replace the exact expression for the error merely by some estimate of its size, and I have considered this for the cases where λ is very small, and where λ is very large.

We can obtain an alternative expression for η_l involving $l(l+1)$ in place of $(l+\frac{1}{2})^2$ in the integrands of (1.3), and with a slightly different error term to compensate. The proofs of the two expressions are essentially the same, except that to prove the second it is not necessary to consider the preliminary transformation of the differential equation (1.1), which is necessary to prove (1.3) [cf. (2) § 6]. I have decided to restrict myself to proving (1.3) since this is the form which is at present more widely recognized, but I have stated (without proof) the corresponding theorem in the alternative case.

2. The expression for the phase shift

Suppose that $q(r)$ satisfies the following conditions:

(i) except at $r = 0$, $q(r)$ is continuously differentiable, while, as $r \to 0$, $q(r) = O(r^{-2+c})$ for some fixed c satisfying $0 < c \leqslant 2$;

(ii) if $V(r) = q(r) + (l+\frac{1}{2})^2/r^2$, then $\lambda - V(r)$ has only one zero, where λ is some positive constant, and this zero is simple;

(iii) $q(r) \to 0$ as $r \to \infty$;

(iv) as $r \to \infty$,

either (*a*) $\int^{\infty} |q(r)|\, dr$ exists

or (*b*) $\int^{\infty} q(r)\, dr$ and $\int^{\infty} |q'(r)|\, dr$ exist.

(It may be remarked that (iv) (*b*) implies (iii), although, of course, (iv) (*a*) does not. For the existence of $\int^{\infty} |q'(r)|\,dr$ implies the existence of $\int^{\infty} q'(r)\,dr$, i.e. that $q(r)$ tends to a limit as $r \to \infty$; and then the existence of $\int^{\infty} q(r)\,dr$ implies that the limit must be zero.)

We then have

THEOREM 1. *Under conditions* (i)–(iv), *we have*

$$\eta_l = \int\limits_{(l+\frac{1}{2})/\surd\lambda}^{\infty} \left\{\lambda - \frac{(l+\frac{1}{2})^2}{r^2}\right\}^{\frac{1}{2}} dr - \int\limits_{r_0}^{\infty} \{\lambda - V(r)\}^{\frac{1}{2}}\,dr + $$

$$+\frac{1}{4}\left\{\pi - \int\limits_{r_0}^{\infty} \frac{d[r^2\{\lambda - V(r)\}]/dr}{r^2\{\lambda - V(r)\}} \sin 2\omega\,dr\right\}, \tag{2.1}$$

where

$$\omega = \omega(r) = \tan^{-1}[\{\lambda - V(r)\}^{\frac{1}{2}} r^{-\frac{1}{2}} Y(r,\lambda)/d\{r^{-\frac{1}{2}} Y(r,\lambda)\}/dr], \tag{2.1 a}$$

and $Y(r,\lambda)$ is the solution of (1.1) *which vanishes at the origin, the existence and uniqueness* (*apart from an unimportant constant factor*) *of $Y(r,\lambda)$ being guaranteed by Lemma* 3 (a) *of* (**2**); *the correct branch of* $\tan^{-1}$ *is determined by insisting that $\omega(r_0) = 0$ and that ω be continuous.*

To obtain the alternative expression for η_l involving integrals with $l(l+1)$ in place of $(l+\frac{1}{2})^2$, suppose that $q(r)$ satisfies the conditions (i), (iii), (iv) as before, but that (ii) is replaced by

(ii)* (*a*) if $l \neq 0$ and if $U(r) = q(r) + l(l+1)/r^2$, then $\lambda - U(r)$ has only one zero, where λ is some positive constant, and this zero is simple;

(*b*) if $l = 0$, so that $U(r) = q(r)$, then $\lambda - U(r)$ has either no zero at all, or else at most one, and that one simple; further, if $\lambda - U(r)$ has no zero, then, as $r \to 0$,

$$U'(r) = O[r^{-2+f}\{\lambda - U(r)\}^{\frac{1}{2}}],$$

for some fixed $f > 0$.

We then have

THEOREM 2. *Under conditions* (i), (ii)*, (iii), (iv), *we have*

$$\eta_l = \int\limits_{\{l(l+1)/\lambda\}^{\frac{1}{2}}}^{\infty} \left\{\lambda - \frac{l(l+1)}{r^2}\right\}^{\frac{1}{2}} - \int\limits_{r_1}^{\infty} \{\lambda - U(r)\}^{\frac{1}{2}}\,dr + $$

$$+\frac{1}{4}\left\{2\pi[\{l(l+1)\}^{\frac{1}{2}} - l] + \int\limits_{r_1}^{\infty} \frac{U'(r)}{\lambda - U(r)} \sin 2\theta\,dr\right\}, \tag{2.2}$$

where r_1 *is the zero of* $\lambda - U(r)$ *(and is taken to be zero if* $\lambda - U(r)$ *has no zero) and*

$$\theta = \theta(r) = \tan^{-1}[\{\lambda - U(r)\}^{\frac{1}{2}} Y(r,\lambda)/Y'(r,\lambda)],$$

the correct branch of $\tan^{-1}$ *being determined by insisting that* $\theta(r_1) = 0$ *and that* θ *be continuous.*

(2.1) is proved in § 3, and in later sections I consider the behaviour of the error term

$$\frac{1}{4}\left\{\pi - \int_{r_0}^{\infty} \frac{d[r^2\{\lambda - V(r)\}]/dr}{r^2\{\lambda - V(r)\}} \sin 2\omega \, dr\right\} \tag{2.3}$$

for large λ and small λ.

Finally, we can see from (2.1), or, of course, alternatively from (2.2), the significance of the term $\frac{1}{2}\pi(l+1)$ in (1.2). For, if $q(r) = 0$, then, from (2.1), η_l is given by (2.3), where now $V(r)$ has the particularly simple form $(l+\frac{1}{2})^2/r^2$. The integral in (2.3) can be evaluated by the method of (**3**). In fact, the evaluation is carried through in (**2**) § 6, for the case where λ is replaced by kr^{-a} $(0 < a < 2)$, k being a positive constant, and the integral is then shown to be π. Both the analysis and the result hold equally well when $a = 0$, so that, in the particular case considered, $\eta_l = 0$. Thus η_l becomes the difference in phase between the physically relevant solution of (1.1) when the potential $q(r)$ is non-existent and the physically relevant solution when $q(r)$ is introduced, leading to the name *phase shift*.

3. The proof of the phase-shift expressions

Under either of the two conditions (*a*), (*b*) given in (iv) of § 2, we know respectively from (**4**) § 5.3 and from (**5**) that, as $r \to \infty$, there are constants A, ϵ (of which ϵ is unique *modulo* π) such that

$$Y(r,\lambda) = A\cos\{r\sqrt{\lambda} - \tfrac{1}{2}\pi(l+1) - \epsilon\} + o(1),$$
$$Y'(r,\lambda) = -A\sqrt{\lambda}\sin\{r\sqrt{\lambda} - \tfrac{1}{2}\pi(l+1) - \epsilon\} + o(1),$$

where $Y(r,\lambda)$ is defined in the statement of Theorem 1. In view of the definition of phase shift, we may replace ϵ by η_l, and then, substituting these asymptotic forms in (2.1 a), we obtain, as $r \to \infty$,

$$\omega(r) = \tfrac{1}{2}\pi + r\sqrt{\lambda} - \tfrac{1}{2}\pi(l+1) - \eta_l \pm m\pi + o(1), \tag{3.1}$$

where m is some positive integer or zero.

At the same time, it is readily verified from (2.1 a) that

$$\frac{d\omega}{dr} = \{\lambda - V(r)\}^{\frac{1}{2}} + \frac{d[r^2\{\lambda - V(r)\}]/dr}{4r^2\{\lambda - V(r)\}} \sin 2\omega.$$

Integrating this over (r_0, R), and using $\omega(r_0) = 0$ and (3.1) with R in

place of r, we obtain that, as $R \to \infty$,

$$R\sqrt{\lambda}-\tfrac{1}{2}\pi l-\eta_l \pm m\pi+o(1)$$

$$= \int_{r_0}^{R} \{\lambda - V(r)\}^{\frac{1}{2}}\, dr + \tfrac{1}{4}\int_{r_0}^{R} \frac{d[r^2\{\lambda-V(r)\}]/dr}{r^2\{\lambda-V(r)\}} \sin 2\omega\, dr. \tag{3.2}$$

This analysis applies in particular when $q(r)$ is identically zero. In this case, we saw in § 2 that $\eta_l = 0$, while the second term on the right-hand side of (3.2) becomes $\frac{1}{4}\pi$, and so (3.2) reduces to

$$R\sqrt{\lambda}-\tfrac{1}{2}\pi l \pm m'\pi + o(1) = \int_{(l+\frac{1}{2})/\sqrt{\lambda}}^{R} \left\{\lambda - \frac{(l+\frac{1}{2})^2}{r^2}\right\}^{\frac{1}{2}} dr + \tfrac{1}{4}\pi, \tag{3.3}$$

where m' is some positive integer or zero.

Subtracting (3.2) from (3.3), and remembering that η_l is determined only *modulo* π, so that we may set $m = m'$, we obtain the required expression (2.1) for the phase shift η_l.

4. In this section it is my purpose to show how to estimate the size of the error term (2.3) for large values of λ. The main result is contained in

THEOREM 3. *If the conditions* (i), (ii), (iii), (iv) *of* § 2 *hold for all sufficiently large* λ, *together with the conditions that*

(A) $q(r)$ *is twice continuously differentiable, except possibly at* $r = 0$;

(B) *as* $r \to 0$, $q'(r) = O(r^{-3+c})$, $q''(r) = O(r^{-4+c})$;

(C) *as* $r \to \infty$, $rq'(r) = O(1)$;

(D) $\int^{\infty} |q''(r)|\, dr$ *exists*;

then the error term (2.3) *is* $O(\lambda^{-\frac{1}{2}c})$, *except in the case* $c = 1$, *when it is* $O(\lambda^{-\frac{1}{2}} \log \lambda)$.

There is a corresponding result for the error term in Theorem 2, i.e.

$$\frac{1}{4}\left\{2\pi[\{l(l+1)\}^{\frac{1}{2}}-l] + \int_{r_1}^{\infty} \frac{U'(r)}{\lambda - U(r)} \sin 2\theta\, dr\right\}. \tag{4.1}$$

I shall state this without proof since the proof is sufficiently similar to that of Theorem 3.

THEOREM 4. *If the conditions* (i), (ii)*, (iii), (iv) *of* § 2 *hold for all sufficiently large* λ, *together with the conditions* (A), (B), (C), (D) *of Theorem* 3, *then the error term* (4.1) *is* $O(\lambda^{-\frac{1}{2}c})$, *except in the case* $c = 1$, *when it is* $O(\lambda^{-\frac{1}{2}} \log \lambda)$.

Two further points should be made. The first is that there are circumstances in which the error term (2.3) or (4.1) is as large as the main term in the expression (2.1) or (2.2) for the phase shift, so that the main term becomes valueless as an approximation to η_l. I discuss this in § 5.

The second point is that the estimates obtained for (2.3) and (4.1) are reached by an application of a method of successive approximations. It would be quite possible to carry the successive approximations further than I have done in order to obtain the precise form of the term in $\lambda^{-\frac{1}{2}c}$ and an estimate of the order of the next largest power of λ involved, and even to go further still, but the process becomes technically rather complicated, and the method is adequately revealed in what I have done.

Proof of Theorem 3. We have already seen in § 2 that, if $q(r) \equiv 0$, then the integral in (2.3) is just π. Hence, without the unimportant factor $\frac{1}{4}$, the error term may be written as

$$\int\limits_{(l+\frac{1}{2})/\sqrt{\lambda}}^{\infty} \frac{d[t^2\{\lambda-(l+\frac{1}{2})^2/t^2\}]/dt}{t^2\{\lambda-(l+\frac{1}{2})^2/t^2\}} \sin 2\omega_0\, dt -$$

$$- \int\limits_{r_0}^{\infty} \frac{d[r^2\{\lambda-V(r)\}]/dr}{r^2\{\lambda-V(r)\}} \sin 2\omega\, dr, \tag{4.2}$$

where $\omega_0 = \omega_0(t)$ is the particular form of ω when $q(r) \equiv 0$.

We now make, in the second integral of (4.2), the change of variable

$$r^2\{\lambda-V(r)\} = t^2\{\lambda-(l+\tfrac{1}{2})^2/t^2\}, \tag{4.3}$$

i.e.

$$t = r\{1-q(r)/\lambda\}^{\frac{1}{2}}. \tag{4.4}$$

This is a valid substitution because

$$\frac{dt}{dr} = \left\{1-\frac{q(r)}{\lambda}\right\}^{\frac{1}{2}} - \frac{\frac{1}{2}rq'(r)/\lambda}{\{1-q(r)/\lambda\}^{\frac{1}{2}}} > 0 \quad \text{for all } r \text{ in } (r_0, \infty), \tag{4.5}$$

provided that λ is sufficiently large, since r_0 is of the order of $\lambda^{-\frac{1}{2}}$ and $q(r)$ satisfies the conditions (i), (iii), (B), (C) of Theorem 3.

Under the change of variable, (4.2) becomes

$$\int\limits_{(l+\frac{1}{2})/\sqrt{\lambda}}^{\infty} \frac{d[t^2\{\lambda-(l+\frac{1}{2})^2/t^2\}]/dt}{t^2\{\lambda-(l+\frac{1}{2})^2/t^2\}} \{\sin 2\omega_0(t) - \sin 2\omega(r)\}\, dt,$$

which, of course, can be written more simply as

$$\int\limits_{(l+\frac{1}{2})/\sqrt{\lambda}}^{\infty} \frac{2\lambda t}{\lambda t^2-(l+\frac{1}{2})^2} \{\sin 2\omega_0(t) - \sin 2\omega(r)\}\, dt. \tag{4.6}$$

To estimate this error term, we must clearly estimate the quantity $\sin 2\omega_0(t) - \sin 2\omega(r)$, and this is a question of estimating the solution $Y(r,\lambda)$ of (1.1) which vanishes at the origin in terms of the corresponding solution $t^{\frac12}J_{l+\frac12}(t\sqrt{\lambda})$ of

$$\frac{d^2y}{dt^2}+\left\{\lambda-\frac{(l+\frac12)^2}{t^2}\right\}y = 0.$$

Such an estimate is contained in Lemmas α, β, and we now interrupt the proof of Theorem 3 to state and prove these lemmas.

LEMMA α. *Under the conditions of Theorem* 3, *where the alternative* (iv) (*a*) *is chosen, both*

$$r^{-\frac12}Y(r,\lambda)-J_{l+\frac12}(t\sqrt{\lambda}) \quad \textit{and} \quad \lambda^{-\frac12}\frac{d}{dr}\{r^{-\frac12}Y(r,\lambda)\}-J'_{l+\frac12}(t\sqrt{\lambda})$$

are of the following forms:

$$\begin{array}{ll}
O(\lambda^{-\frac12 c}) & \textit{for} \quad (l+\tfrac12)\lambda^{-\frac12} \leqslant t \leqslant K\lambda^{-\frac12}, \\
 & \quad K \textit{ being any constant exceeding } l+\tfrac12; \\
O(t^{-\frac12}\lambda^{-\frac14-\frac12 c}) & \textit{for} \quad c<1 \quad \textit{and} \quad K\lambda^{-\frac12} \leqslant t \leqslant a, \\
 & \quad a \textit{ being any positive constant}; \\
O(t^{c-\frac12}\lambda^{-\frac34}) & \textit{for} \quad c>1 \quad \textit{and} \quad K\lambda^{-\frac12} \leqslant t \leqslant a; \\
O(t^{-\frac12}\lambda^{-\frac34}\log\lambda) & \textit{for} \quad c=1 \quad \textit{and} \quad K\lambda^{-\frac12} \leqslant t \leqslant a; \\
O(t^{-\frac12}\lambda^{-\frac14-\frac12 c}) & \textit{for} \quad c<1 \quad \textit{and} \quad a \leqslant t < \infty; \\
O(t^{\frac12}\lambda^{-\frac34}) & \textit{for} \quad c>1 \quad \textit{and} \quad a \leqslant t < \infty; \\
O(t^{-\frac12}\lambda^{-\frac34}\log\lambda)+O(t^{\frac12}\lambda^{-\frac34}) & \textit{for} \quad c=1 \quad \textit{and} \quad a \leqslant t < \infty.
\end{array}$$

Proof. The proof is a modification of the argument contained in § 3 of (6). $Y(r,\lambda)$ satisfies (1.1) and also the condition $Y(0,\lambda)=0$, and so $Y(r,\lambda)$ satisfies the integral equation

$$Y(r,\lambda) = r^{\frac12}J_{l+\frac12}(r\sqrt{\lambda})+\tfrac12\pi(-1)^{l+1}\int\limits_0^r \{J_{l+\frac12}(r\sqrt{\lambda})J_{-l-\frac12}(u\sqrt{\lambda})-$$
$$-J_{l+\frac12}(u\sqrt{\lambda})J_{-l-\frac12}(r\sqrt{\lambda})\}r^{\frac12}u^{\frac12}q(u)Y(u,\lambda)\,du. \qquad (4.7)$$

Let
$$v(r,\lambda) = \begin{cases} r^{l+1}\lambda^{\frac12 l+\frac14} & (r \leqslant k\lambda^{-\frac12}), \\ \lambda^{-\frac14} & (r > k\lambda^{-\frac12}), \end{cases}$$

where k is some fixed positive constant. Then

$$|r^{\frac12}J_{l+\frac12}(r\sqrt{\lambda})| < Av(r,\lambda),$$

where A denotes various positive constants. Let

$$\max_{0\leqslant r<\infty}\frac{|Y(r,\lambda)|}{v(r,\lambda)} = M.$$

Then, if $r \leqslant k\lambda^{-\frac{1}{2}}$, (4.7) gives

$$M < A + AM \int\limits_0^r \left(\frac{r}{u}\right)^{l+\frac{1}{2}} r^{\frac{1}{2}} u^{\frac{1}{2}} |q(u)| \frac{v(u,\lambda)}{v(r,\lambda)}\, du < A + AMr^c,$$

so that, for $r \leqslant k\lambda^{-\frac{1}{2}}$, $\quad |Y(r,\lambda)| < Av(r,\lambda)$.

Substituting this back in (4.7), we obtain

$$Y(r,\lambda) = r^{\frac{1}{2}} J_{l+\frac{1}{2}}(r\surd\lambda) + O(r^{l+1+c}\lambda^{\frac{1}{2}l+\frac{1}{4}}),$$

i.e.
$$r^{-\frac{1}{2}} Y(r,\lambda) = J_{l+\frac{1}{2}}(r\surd\lambda) + O(r^{l+\frac{1}{2}+c}\lambda^{\frac{1}{2}l+\frac{1}{4}}).$$

Now from the definition (4.4) of t in terms of r, it is clear that over the complete range $\{(l+\frac{1}{2})\lambda^{-\frac{1}{2}}, \infty\}$ for t, t and r are of the same order, and so, in the range $\{(l+\frac{1}{2})\lambda^{-\frac{1}{2}}, K\lambda^{-\frac{1}{2}}\}$ for t, r is of the order $\lambda^{-\frac{1}{2}}$, and

$$t = r[1+O\{q(r)/\lambda\}] = r\{1+O(\lambda^{-\frac{1}{2}c})\}.$$

Hence

$$r^{-\frac{1}{2}} Y(r,\lambda) = J_{l+\frac{1}{2}}[t\surd\lambda\{1+O(\lambda^{-\frac{1}{2}c})\}] + O(\lambda^{-\frac{1}{2}c}) = J_{l+\frac{1}{2}}(t\surd\lambda) + O(\lambda^{-\frac{1}{2}c}),$$

as required.

For the range $\{K\lambda^{-\frac{1}{2}}, a\}$ of t, we shall consider only the case $c > 1$. The other cases are sufficiently similar to require no separate mention. Then, if $k\lambda^{-\frac{1}{2}} \leqslant r \leqslant b$, for some positive constant b, (4.7) gives

$$M < A + AM \int\limits_0^{k\lambda^{-\frac{1}{2}}} u|q(u)|\, du + AM\lambda^{-\frac{1}{2}} \int\limits_{k\lambda^{-\frac{1}{2}}}^{r} |q(u)|\, du < A + AMr^{c-1}\lambda^{-\frac{1}{2}},$$

so that $|Y(r,\lambda)| < Av(r,\lambda)$. Substituting this back in (4.7), we obtain

$$Y(r,\lambda) = r^{\frac{1}{2}} J_{l+\frac{1}{2}}(r\surd\lambda) + O(r^{c-1}\lambda^{-\frac{3}{4}}). \tag{4.8}$$

Now setting
$$t = r\{1+O(r^{-2+c}/\lambda)\}, \tag{4.9}$$

i.e.
$$r = t\{1+O(t^{-2+c}/\lambda)\}, \tag{4.10}$$

we obtain the required estimate in the range $\{K\lambda^{-\frac{1}{2}}, a\}$ of t.

The final range $\{a, \infty\}$ of t is dealt with similarly.

To obtain the estimates for $d\{r^{-\frac{1}{2}}Y(r,\lambda)\}/dr$ given in the statement of the lemma, we multiply (4.7) by $r^{-\frac{1}{2}}$, differentiate with respect to r, and work with the resulting equation in the same way as we worked with (4.7) itself.

This completes the proof of Lemma α, but we require similar formulae in the case where alternative (iv) (b) is chosen in place of (iv) (a). This is covered by

LEMMA β. *Under the conditions of Theorem* 3, *where the alternative* (iv) (b) *is chosen, we have the same estimates as were obtained in Lemma* α.

Proof. The choice of (iv) (*b*) in place of (iv) (*a*) affects only the range (b,∞) of r. In fact, at $r = b$, we have (4.8) still true. (Again we sufficiently consider only the case $c > 1$.)

Now consider the function, for $r \geqslant b$,

$$z_1(r) = \exp\left\{i\int\limits_b^r \{\lambda - V(u)\}^{\frac{1}{2}}\,du\right\}, \tag{4.11}$$

where, as before, $\quad V(u) = q(u)+(l+\tfrac{1}{2})^2/u^2. \quad$ (4.12)

It is easy to verify that z_1 satisfies the differential equation

$$\frac{d^2z}{dr^2}+\left\{\lambda - V(r)+\tfrac{1}{2}i\frac{V'(r)}{\{\lambda-V(r)\}^{\frac{1}{2}}}\right\}z = 0, \tag{4.13}$$

of which another solution is

$$z_2(r) = z_1(r)\int\limits_b^r z_1^{-2}(u)\,du.$$

Furthermore, the Wronskian $W(z_1,z_2) = 1$, while, for large λ,

$$z_1(r) = \exp[i\{(r-b)\surd\lambda+O(\lambda^{-\frac{1}{2}})\}] = e^{i(r-b)\surd\lambda}\{1+O(\lambda^{-\frac{1}{2}})\},$$

and so $\quad z_2(r) = \lambda^{-\frac{1}{2}}\sin\{(r-b)\surd\lambda\}\{1+O(\lambda^{-\frac{1}{2}})\}.$

We can now write the equation for $Y(r,\lambda)$ in the form

$$\frac{d^2Y}{dr^2}+\left\{\lambda-V(r)+\tfrac{1}{2}i\frac{V'(r)}{\{\lambda-V(r)\}^{\frac{1}{2}}}\right\}Y = \tfrac{1}{2}i\frac{V'(r)}{\{\lambda-V(r)\}^{\frac{1}{2}}}Y,$$

which becomes, as an integral equation,

$$Y(r,\lambda) = Az_1(r)+Bz_2(r)-\int\limits_b^r \{z_1(r)z_2(u)-z_1(u)z_2(r)\}\tfrac{1}{2}i\frac{V'(u)}{\{\lambda-V(u)\}^{\frac{1}{2}}}Y(u,\lambda)\,du,$$

where A, B are suitable constants, dependent on λ, but not, of course, on r. If we solve this by iteration (as, for example, in the proof of Lemma α), we have

$$Y(r,\lambda) = Az_1(r)+Bz_2(r)+O[\lambda^{-1}\{|A|+|B|\lambda^{-\frac{1}{2}}\}].$$

Replacing $z_1(r)$, $z_2(r)$ by their approximate expressions as trigonometric functions, we have, for constants C, D,

$$Y(r,\lambda) = C\cos\{r\surd\lambda-\tfrac{1}{2}(l+\tfrac{1}{2})\pi-\tfrac{1}{4}\pi\}+ \\ +D\sin\{r\surd\lambda-\tfrac{1}{2}(l+\tfrac{1}{2})\pi-\tfrac{1}{4}\pi\}+O\{\lambda^{-\frac{1}{2}}(|C|+|D|)\}. \tag{4.14}$$

This holds for $r \geqslant b$, and so for any fixed range $\{b_1, b_2\}$.

However, from (4.8), we have in the same range $\{b_1, b_2\}$, using the

asymptotic expansions for Bessel functions, that

$$Y(r,\lambda) = \left(\frac{2}{\pi}\right)^{\frac{1}{2}} \lambda^{-\frac{1}{4}} \cos\{r\sqrt{\lambda} - \tfrac{1}{2}(l+\tfrac{1}{2})\pi - \tfrac{1}{4}\pi\} + O(\lambda^{-\frac{3}{4}}). \tag{4.15}$$

Since (4.14) and (4.15) must represent the same function throughout the range $\{b_1, b_2\}$, we conclude that

$$C = \left(\frac{2}{\pi}\right)^{\frac{1}{2}} \lambda^{-\frac{1}{4}}\{1+O(\lambda^{-\frac{1}{2}})\}, \qquad D = O(\lambda^{-\frac{3}{4}}).$$

Hence we may substitute for C, D in (4.14), and so obtain (4.15), holding now, however, for all $r \geqslant b$ and not merely for the range $\{b_1, b_2\}$. Using the asymptotic expansions for Bessel functions, we can rewrite (4.15) as

$$Y(r,\lambda) = r^{\frac{1}{2}} J_{l+\frac{1}{2}}(r\sqrt{\lambda}) + O(\lambda^{-\frac{3}{4}}). \tag{4.15 a}$$

Finally making the transformation from r to t, we find ourselves with the estimate we require.

We can now return to the proof of Theorem 3. We saw in (4.6) that the error term could be written as

$$\int\limits_{(l+\frac{1}{2})/\sqrt{\lambda}}^{\infty} \frac{2\lambda t}{\lambda t^2 - (l+\frac{1}{2})^2} \{\sin 2\omega_0(t) - \sin 2\omega(r)\}\, dt. \tag{4.16}$$

We consider first the contribution to this integral from the range $\{(l+\frac{1}{2})\lambda^{-\frac{1}{2}}, K\lambda^{-\frac{1}{2}}\}$ of t. Then

$$\sin 2\omega_0(t) = \frac{2\{\lambda-(l+\frac{1}{2})^2/t^2\}^{\frac{1}{2}} J_{l+\frac{1}{2}}(t\sqrt{\lambda})\, d\{J_{l+\frac{1}{2}}(t\sqrt{\lambda})\}/dt}{[(d/dt)\{J_{l+\frac{1}{2}}(t\sqrt{\lambda})\}]^2 + \{\lambda-(l+\frac{1}{2})^2/t^2\} J^2_{l+\frac{1}{2}}(t\sqrt{\lambda})},$$

$$\sin 2\omega(r) = \frac{2\{\lambda - V(r)\}^{\frac{1}{2}} r^{-\frac{1}{2}} Y(r,\lambda)\, d\{r^{-\frac{1}{2}} Y(r,\lambda)\}/dr}{[(d/dr)\{r^{-\frac{1}{2}} Y(r,\lambda)\}]^2 + \{\lambda - V(r)\} r^{-1} Y^2(r,\lambda)}.$$

Using the results of Lemmas α, β, and the fact that from (4.4)

$$\{\lambda - V(r)\}^{\frac{1}{2}} = \{\lambda-(l+\tfrac{1}{2})^2/t^2\}^{\frac{1}{2}} t/r = \{\lambda-(l+\tfrac{1}{2})^2/t^2\}\{1+O(\lambda^{-\frac{1}{2}c})\},$$

we have $$\sin 2\omega_0(t) - \sin 2\omega(r) = O[\{\lambda-(l+\tfrac{1}{2})^2/t^2\}^{\frac{1}{2}} \lambda^{-\frac{1}{2}-\frac{1}{2}c}],$$
and so

$$\int\limits_{(l+\frac{1}{2})/\sqrt{\lambda}}^{K/\sqrt{\lambda}} \frac{2\lambda t}{\lambda t^2 - (l+\frac{1}{2})^2} \{\sin 2\omega_0(t) - \sin 2\omega(r)\}\, dt$$

$$= O\left\{\lambda^{-\frac{1}{2}c} \int\limits_{(l+\frac{1}{2})/\sqrt{\lambda}}^{K/\sqrt{\lambda}} \{\lambda-(l+\tfrac{1}{2})^2/t^2\}^{\frac{1}{2}}\, dt\right\} = O(\lambda^{-\frac{1}{2}c}).$$

To consider the contribution to the integral (4.16) from the range $\{K\lambda^{-\frac{1}{2}}, \infty\}$, we split the integral back into separate integrals with respect

to t and r, obtaining

$$\int\limits_{K/\sqrt{\lambda}}^{\infty} \frac{2\lambda t}{\lambda t^2-(l+\frac{1}{2})^2} \sin 2\omega_0(t)\,dt - \int\limits_{K'/\sqrt{\lambda}}^{\infty} \frac{d[r^2\{\lambda-V(r)\}]/dr}{r^2\{\lambda-V(r)\}} \sin 2\omega(r)\,dr, \tag{4.17}$$

where the value $K'/\sqrt{\lambda}$ of r corresponds to the value $K/\sqrt{\lambda}$ of t. Taking the second integral first, we write it as

$$\int\limits_{K'/\sqrt{\lambda}}^{\infty} \frac{d[r^2\{\lambda-V(r)\}]/dr}{r^2\{\lambda-V(r)\}} \frac{\sin 2\omega(d\omega/dr)}{\{\lambda-V(r)\}^{\frac{1}{2}}}\,dr - \\ -\frac{1}{4}\int\limits_{K'/\sqrt{\lambda}}^{\infty} \left\{\frac{d[r^2\{\lambda-V(r)\}]/dr}{r^2\{\lambda-V(r)\}}\right\}^2 \frac{\sin^2 2\omega}{\{\lambda-V(r)\}^{\frac{1}{2}}}\,dr, \tag{4.18}$$

using the expression for $d\omega/dr$ obtained in § 3. The first integral in (4.18) can now be integrated by parts, and (4.18) becomes

$$\left[-\frac{1}{2}\frac{d[r^2\{\lambda-V(r)\}]/dr}{r^2\{\lambda-V(r)\}} \frac{\cos 2\omega}{\{\lambda-V(r)\}^{\frac{1}{2}}}\right]_{K'/\sqrt{\lambda}}^{\infty} + \\ +\frac{1}{2}\int\limits_{K'/\sqrt{\lambda}}^{\infty} \cos 2\omega \frac{d}{dr}\left\{\frac{d[r^2\{\lambda-V(r)\}]/dr}{r^2\{\lambda-V(r)\}^{\frac{1}{2}}}\right\}\,dr - \\ -\frac{1}{4}\int\limits_{K'/\sqrt{\lambda}}^{\infty} \left\{\frac{d[r^2\{\lambda-V(r)\}]/dr}{r^2\{\lambda-V(r)\}}\right\}^2 \frac{\sin^2 2\omega}{\{\lambda-V(r)\}^{\frac{1}{2}}}\,dr.$$

A similar procedure can be carried out on the first integral in (4.17). If we subtract the two results, and consider first the integrated terms, these vanish in both cases at ∞, while at the lower end of integration,

$$[\lambda t^2-(l+\tfrac{1}{2})^2]_{K/\sqrt{\lambda}} = [r^2\{\lambda-V(r)\}]_{K'/\sqrt{\lambda}},$$

$$[2\lambda t]_{K/\sqrt{\lambda}} = \left[\frac{d}{dr}[r^2\{\lambda-V(r)\}]\frac{dr}{dt}\right]_{K'/\sqrt{\lambda}}$$

$$= \left[\frac{d}{dr}[r^2\{\lambda-V(r)\}]\right]_{K'/\sqrt{\lambda}}\{1+O(\lambda^{-\frac{1}{2}c}\},$$

$$[\cos 2\omega_0(t)]_{K/\sqrt{\lambda}} = [\cos 2\omega(r)]_{K'/\sqrt{\lambda}}+O(\lambda^{-\frac{1}{2}c}),$$

the last line of formula being obtained similarly to the corresponding result for $\sin 2\omega$ which was used earlier. Hence the difference between the integrated terms is $O(\lambda^{-\frac{1}{2}c})$.

Turning now to the integrals, in the integrals with respect to r we can make the change of variable to t, so that (considering sufficiently

the second of the two integrals) the difference is

$$-\frac{1}{4}\int\limits_{K/\sqrt{\lambda}}^{\infty}\left[\frac{2\lambda t}{\lambda t^2-(l+\frac{1}{2})^2}\right]^2\left[\frac{\sin^2 2\omega_0(t)}{\{\lambda-(l+\frac{1}{2})^2/t^2\}^{\frac{1}{2}}}-\frac{\sin^2 2\omega(r)}{\{\lambda-V(r)\}^{\frac{1}{2}}}\frac{dt}{dr}\right]dt.$$

We break this integral into the ranges $\{K/\sqrt{\lambda}, a\}$, $\{a, \infty\}$, and we shall consider sufficiently $c < 1$.

Then, in $\{K/\sqrt{\lambda}, a\}$,

$$\{\lambda-V(r)\}^{\frac{1}{2}} = \{\lambda-(l+\tfrac{1}{2})^2/t^2\}^{\frac{1}{2}}\{1+O(t^{-2+c}/\lambda)\},$$

$$\frac{dt}{dr} = 1+O(t^{-2+c}/\lambda),$$

$$\sin^2 2\omega_0(t)-\sin^2 2\omega(r) = O(t^{-1}\lambda^{-\frac{1}{2}-\frac{1}{2}c}).$$

Hence the contribution to the integral is

$$O\left\{\int\limits_{K/\sqrt{\lambda}}^{a}\frac{1}{\lambda^{\frac{1}{2}}t^2}(t^{-1}\lambda^{-\frac{1}{2}-\frac{1}{2}c}+t^{-2+c}\lambda^{-1})\,dt\right\},$$

and, since $t^{-2+c}\lambda^{-1}$ can be neglected in comparison with $t^{-1}\lambda^{-\frac{1}{2}-\frac{1}{2}c}$ for $t \geqslant K/\sqrt{\lambda}$ and $c \leqslant 1$, the contribution is

$$O\left(\int\limits_{K/\sqrt{\lambda}}^{a} t^{-3}\lambda^{-1-\frac{1}{2}c}\,dt\right) = O(\lambda^{-\frac{1}{2}c}).$$

In $\{a, \infty\}$,

$$\{\lambda-V(r)\}^{\frac{1}{2}} = \{\lambda-(l+\tfrac{1}{2})^2/t^2\}^{\frac{1}{2}}\{1+O(\lambda^{-1})\},$$

$$\frac{dt}{dr} = 1+O(\lambda^{-1}),$$

$$\sin^2 2\omega_0(t)-\sin^2 2\omega(r) = O(t^{-1}\lambda^{-\frac{1}{2}-\frac{1}{2}c})+O(\lambda^{-1}).$$

Hence the contribution to the integral is

$$O\left\{\int\limits_{a}^{\infty}\frac{1}{\lambda^{\frac{1}{2}}t^2}(t^{-1}\lambda^{-\frac{1}{2}-\frac{1}{2}c}+\lambda^{-1})\,dt\right\} = O(\lambda^{-1-\frac{1}{2}c}+\lambda^{-\frac{3}{2}}) = O(\lambda^{-1-\frac{1}{2}c}).$$

This completes the proof of Theorem 3.

5. I return now to the point which I mentioned in stating Theorems 3 and 4: that the error terms (2.3) or (4.1) may be as large as the main term in the expressions (2.1) or (2.2) for the phase shift, so that the main term becomes rather valueless as an approximation to η_l.

Suppose that the conditions of Theorem 3 (or Theorem 4) are satisfied, and consider $c < 1$. In this case, the formula corresponding to (4.15 a) (which was developed for $c > 1$) has an error term $O(\lambda^{-\frac{1}{2}-\frac{1}{2}c})$ in place of the error term $O(\lambda^{-1})$, as may be readily verified by carrying through

the analysis. (In fact, (4.15 a) was obtained under the conditions of Lemma β, but the same formula would have appeared under Lemma α had we written out the full details of the proof.) Hence, as $r \to \infty$,

$$\begin{aligned} r^{-\frac{1}{2}}Y(r,\lambda) &= J_{l+\frac{1}{2}}(r\sqrt{\lambda})+O(r^{-\frac{1}{2}}\lambda^{-\frac{1}{4}-\frac{1}{2}c}) \\ &= \left(\frac{2}{\pi r\sqrt{\lambda}}\right)^{\frac{1}{2}}\cos\{r\sqrt{\lambda}-\tfrac{1}{2}(l+\tfrac{1}{2})\pi-\tfrac{1}{4}\pi\}+O(r^{-\frac{1}{2}}\lambda^{-\frac{1}{4}-\frac{1}{2}c}), \end{aligned}$$

and

$$\lambda^{-\frac{1}{2}}\frac{d}{dr}\{r^{-\frac{1}{2}}Y(r,\lambda)\} = -\left(\frac{2}{\pi r\sqrt{\lambda}}\right)^{\frac{1}{2}}\sin\{r\sqrt{\lambda}-\tfrac{1}{2}(l+\tfrac{1}{2})\pi-\tfrac{1}{4}\pi\}+O(r^{-\frac{1}{2}}\lambda^{-\frac{1}{4}-\frac{1}{2}c}).$$

Substituting this in (2.1 a), we obtain, as $r \to \infty$,

$$\omega(r) = \tfrac{1}{2}\pi+r\sqrt{\lambda}-\tfrac{1}{2}(l+\tfrac{1}{2})\pi-\tfrac{1}{4}\pi+O(\lambda^{-\frac{1}{2}c})\pm m'\pi+o(1),$$

where m' is some positive integer or zero.

However, we also have (3.1), and comparing this with the last equation, we obtain

$$\eta_l = \pm(m-m')\pi+O(\lambda^{-\frac{1}{2}c}).$$

Since we can deduce fairly readily from (2.1) that η_l is small for large λ, we put $m = m'$ and obtain

$$\eta_l = O(\lambda^{-\frac{1}{2}c}).$$

On the other hand, all that we were able to say of the error term in § 4 was that it too was $O(\lambda^{-\frac{1}{2}c})$, and, unless this estimate of the error is an unnecessarily large one, the error term is as large as η_l itself.

In fact, the estimate $O(\lambda^{-\frac{1}{2}c})$ for the error term is in general best-possible. I do not intend to give a detailed proof of this, but it can be shown straightforwardly enough, though tediously, by carrying through the argument of § 4 more carefully and noting that the 'coefficient' of $\lambda^{-\frac{1}{2}c}$ in the error term does not in general vanish.

If we take $c = 1$ instead of $c < 1$, the same considerations apply and the error term may be as large as η_l itself. This does not, however, apply for $c > 1$. For then the argument above yields $\eta_l = O(\lambda^{-\frac{1}{2}})$, and in fact, in general, $\eta_l \asymp \lambda^{-\frac{1}{2}}$, while the error term, as proved in § 4, is $O(\lambda^{-\frac{1}{2}c})$, which is now of lower order.

6. In this section it is my purpose to show how to estimate the size of the error term (2.3) for small values of λ. The main result is contained in the theorem:

THEOREM 5. *Suppose that conditions* (i), (ii), (iii), (iv) *of* § 2 *hold for all sufficiently small* λ, *together with the conditions*

(α) $V(r)$ *has one and only one zero, and that one simple, from which it follows that* $V(r) < 0$, $q(r) < 0$ *for sufficiently large* r;

(β) *as* $r \to \infty$, $V(r)$ *is twice continuously differentiable, with*

$$\frac{V'(r)}{-V(r)} \asymp \frac{1}{r}, \qquad \frac{V''(r)}{V'(r)} = O\left(\frac{1}{r}\right);$$

(γ) *as* $r \to \infty$, $r^{d-2}\{-V(r)\}^{-1}$ *is a decreasing function for some fixed* $d > 0$.

Then, if $\omega^*(r)$ *is the particular form of* $\omega(r)$ *when* $\lambda = 0$, *we have*

$$\int\limits_{r_0}^{\infty} \frac{d[r^2\{\lambda - V(r)\}]/dr}{r^2\{\lambda - V(r)\}} \sin 2\omega \, dr$$

$$= \int\limits_{r_0^*}^{\infty} \frac{d\{r^2 V(r)\}/dr}{r^2 V(r)} \sin 2\omega^* \, dr + O\{\lambda^{\frac{1}{2}} V^{-\frac{1}{2}}(\lambda^{-\frac{1}{2}})\}, \tag{6.1}$$

where r_0^* *is the zero of* $V(r)$.

This means that the error term (2.3) becomes approximately

$$\frac{1}{4}\left[\pi - \int\limits_{r_0^*}^{\infty} \frac{d\{r^2 V(r)\}/dr}{r^2 V(r)} \sin 2\omega^* \, dr\right], \tag{6.2}$$

though there is no necessity for this to be zero. However, it is curious that, if $q(r)$ takes the special form $-kr^{-a}$ $(1 < a < 2)$, k being some positive constant, then the expression (6.2) is indeed zero. This is proved in (**2**), at the end of § 6.

Theorem 5 is in effect proved in the course of (**2**). However, what is more directly proved is the corresponding result for the error term of Theorem 2, i.e. (4.1), and I shall state this now and then indicate where in (**2**) the proof of this (and so sufficiently the proof of Theorem 5) can be found.

Theorem 6. *Suppose that the conditions* (i), (ii)*, (iii), (iv) *of* § 2 *hold for all sufficiently small* λ, *together with the conditions*

(α)* (*a*) *if* $l \neq 0$, $U(r)$ *has one and only one zero, and that one simple, from which it follows that* $U(r) < 0$, $q(r) < 0$ *for sufficiently large* r;

(*b*) *if* $l = 0$, *so that* $U(r) = q(r)$, *then* $U(r)$ *has at most one zero, and that one simple, and* $U(r) < 0$ *for sufficiently large* r; *further, if* $U(r)$ *has no zero, then, as* $r \to 0$,

$$U'(r) = O\{r^{-2+h} U^{\frac{1}{2}}(r)\}$$

for some fixed $h > 0$;

(β)* *as* $r \to \infty$, $U(r)$ *is twice continuously differentiable, with*

$$\frac{U'(r)}{-U(r)} \asymp \frac{1}{r}, \qquad \frac{U''(r)}{U'(r)} = O\left(\frac{1}{r}\right);$$

(γ)* *as* $r \to \infty$, $r^{d-2}\{-U(r)\}^{-1}$ *is a decreasing function for some fixed* $d > 0$.

Then, if $\theta^*(r)$ *is the particular form of* $\theta(r)$ *when* $\lambda = 0$, *we have*

$$\int_{r_1}^{\infty} \frac{U'(r)}{\lambda - U(r)} \sin 2\theta \, dr = -\int_{r_1^*}^{\infty} \frac{U'(r)}{U(r)} \sin 2\theta^* \, dr + O\{\lambda^{\frac{1}{2}} U^{-\frac{1}{2}}(\lambda^{-\frac{1}{2}})\}, \tag{6.3}$$

where r_1^* *is the zero of* $U(r)$, *if this exists, and is otherwise zero.*

This means that the error term (4.1) becomes approximately

$$\frac{1}{4}\left\{2\pi[\{l(l+1)\}^{\frac{1}{2}} - l] - \int_{r_1^*}^{\infty} \frac{U'(r)}{U(r)} \sin 2\theta^* \, dr\right\}, \tag{6.4}$$

though again this is not necessarily zero. In fact, if $q(r)$ takes the special form $-kr^{-a}$ $(1 < a < 2)$, k being some positive constant, then from [(2) equation (2.3)], (6.4) becomes

$$\frac{1}{4}\left[\frac{\pi a}{2-a}[2l+1-2\{l(l+1)\}^{\frac{1}{2}}]\right].$$

To prove Theorem 6, we consider the expression

$$\int_{r_1}^{\infty} \frac{U'(r)}{\lambda - U(r)} \sin 2\theta \, dr + \int_{r_1^*}^{\infty} \frac{U'(r)}{U(r)} \sin 2\theta^* \, dr,$$

and that this is indeed $O\{\lambda^{\frac{1}{2}} U^{-\frac{1}{2}}(\lambda^{-\frac{1}{2}})\}$ is proved in [(2) Lemmas 5 (a)–5 (d)] together with the result, proved similarly to Lemma 5 (d), that

$$\int_{\lambda^{-\frac{1}{2}}}^{\infty} \frac{U'(r)}{\lambda - U(r)} \sin 2\theta \, dr = O\{\lambda^{\frac{1}{2}} U^{-\frac{1}{2}}(\lambda^{-\frac{1}{2}})\}.$$

Negative λ_n has to be replaced by positive λ throughout, but this is an unimportant alteration, and it should be remarked too that the conditions imposed in (2) are slightly more restrictive than we are employing here. It will be readily verified by the reader, however, that the extra restrictions in (2) are required, not for the lemmas to which we appeal here, but for Lemmas 5 (e), 5 (f).

REFERENCES

1. P. Morse and H. Feshbach, *Methods of theoretical physics*, vol. 2 (New York, 1953).
2. J. B. McLeod, 'The distribution of the eigenvalues for the hydrogen atom and similar cases', *Proc. London Math. Soc.* (3) 11 (1961) 139–58.

3. J. B. McLeod, 'On certain integral formulae', *Proc. London Math. Soc.* (3) 11 (1961) 134–8.

4. E. C. Titchmarsh, *Eigenfunction expansions associated with second-order differential equations* (Oxford, 1946).

5. A. Wintner, 'Asymptotic integration of the adiabatic oscillator', *American J. of Math.* 69 (1947) 251–72.

6. E. C. Titchmarsh, 'On the eigenvalues in problems with spherical symmetry', *Proc. Royal Soc.* A 245 (1958) 147–55.

RIESZ SUMMABILITY OF SUBSEQUENCES

By F. R. KEOGH and G. M. PETERSEN (Swansea)

[Received 17 February 1960]

SUPPOSE that $\{s_n\}$ is a real sequence and t is a real number in the interval $0 \leqslant t \leqslant 1$. Representing t by a non-terminating binary decimal expansion, we shall denote by $\{s_n(t)\}$ the subsequence obtained from $\{s_n\}$ by omitting s_k if and only if there is a 0 in the kth decimal place in the expansion of t. With this correspondence, it is then possible to speak of 'a set of subsequences of measure one', 'an everywhere dense set of subsequences', and so on. We shall alternatively denote the subsequence $\{s_n(t)\}$ by $\{s_{k_n}\}$, where such notation is more useful. We adopt the convention that where no limits are stated in the summation sign, sums are to be taken from 0 to ∞ (or 1 to ∞ in some evident cases).

A theorem by Buck and Pollard (**3**), is stated as follows:

THEOREM 1. *If $\{s_n\}$ is $(C,1)$ summable to S and*

$$\sum s_k^2/k^2 < \infty,$$

then almost all the subsequences are $(C,1)$ summable to S.

This theorem has been extended in one direction by Hill (**7**). We wish to extend it for bounded sequences to Riesz means of a restricted type. First, however, we state three lemmas which will be of use in the proof; they may be found in (**3**). The properties of the Rademacher functions $R_n(t)$ can be found in (**8**).

LEMMA 1. *Let $\{P_n\}$ be a sequence of positive numbers, increasing monotonically to infinity. If the series $\sum a_n P_n$ converges, then*

$$\lim_{n\to\infty} \frac{1}{P_n} \sum_1^n a_k = 0.$$

LEMMA 2. *The series $\sum a_k R_k(t)$ converges on a set of measure one or measure zero according as $\sum a_k^2$ converges or not.*

Define the regular Riesz means (R, p_k) by

$$\lim_{n\to\infty} \frac{p_0 s_0 + \dots + p_n s_n}{P_n},$$

where $P_n = p_0 + \dots + p_n$.

LEMMA 3. *For the regular Riesz means, if* $\sum (p_n/P_n)^2$ *converges and* $\{s_k\}$ *is bounded,*

$$\lim_{n\to\infty} \frac{1}{P_n} \sum_0^n p_k s_k R_k(t) = 0$$

almost everywhere.

Lemma 3 is proved by an application of Lemma 2 and then Lemma 1. We now prove the lemma:

LEMMA 4. *If* (R, p_n) *is a regular Riesz means,*

$$p_k \leqslant p_{k+1} \quad (k = 0, 1, 2, \ldots),$$

and

$$\frac{P_k}{p_k} \leqslant L\frac{P_n}{p_n} \quad (n = 0, 1, 2, \ldots;\ k = 0, 1, 2, \ldots, n),$$

then

$$\sum_{k=0}^{n-1} \frac{P_k}{p_k} \frac{|p_k - p_{k+1}|}{P_n} \leqslant M.$$

Proof. We have

$$\sum_{k=0}^{n-1} \frac{P_k}{p_k} \frac{|p_k - p_{k+1}|}{P_n} \leqslant L\frac{P_n}{p_n} \sum_{k=0}^{n-1} \frac{|p_k - p_{k+1}|}{P_n} \leqslant L\frac{P_n}{p_n} \frac{p_n - p_0}{P_n},$$

and our assertion is proved.

We also state a theorem due to Borel. Here a sequence $\{s_n\}$ of 0's and 1's corresponds with the number $\alpha = \cdot a_1 a_2 \ldots$ in the non-terminating binary expansion if and only if $a_n = s_n$ for every n. For full details see Hill (**8**) or Cooke (**4**).

THEOREM 2. *Almost all sequences of* 0*'s and* 1*'s are* $(C, 1)$ *summable to* $\frac{1}{2}$.

This theorem means that, if $\alpha(n)$ is the number of 1's in the first n places of α, $\alpha(n) = \frac{1}{2}n + o(n)$. Hence, when we are investigating subsequences and wish to extend Theorem 1, we easily see that the subsequences $\{s_{k_n}\}$ such that $k_n \leqslant 3n$ for $n > N$, correspond to a set of measure one. This means that we can restrict our considerations to subsequences such that $k_n \leqslant 3n$ for $n > N$ since the measure of the intersection of this set with a second set will also be the measure of the second set. Of course, Theorem 2 is not itself concerned with subsequences and is introduced here only for the purpose of restricting the k_n in the manner described.

Before proceeding to our main theorem, we want to provide a counterexample to show that, even for bounded sequences, Theorem 1 cannot be extended by some such statement as:

If $A = (a_{m,n})$ *is stronger than* $(C, 1)$ *then almost all subsequences of a bounded* A*-summable sequence are* A*-summable.*

Consider the sequence $\{m(n)\}$ $(n = 1, 2, 3, \ldots)$, where

$$m(n) = [3m(n-1)]^2, \qquad m(1) = 1,$$

i.e.

$$m(n) = 3^{2^n-2},$$

and the first three terms of the sequence are 1, 9, 729,.... Let

$$a_{n,k} = \frac{1}{m(n)},$$

where

$$3m(n-1) \leqslant k \leqslant m(n), \qquad a_{n,k} = 0$$

for other values of k. The matrix $A = (a_{n,k})$ is regular and, since

$$\sum k|a_{n,k} - a_{n,k+1}| \leqslant M,$$

A is stronger than the Cesàro matrix [see (**12**)]. Let the sequence $\{s_k\}$ be defined as follows:

$$s_k = 1 \quad (m(2r) \leqslant k < 3m(2r)), \qquad s_k = 0 \quad \text{otherwise.}$$

Now $\{s_k\}$ is A-summable to 0. From Theorem 2, however, it is clear that, if we take the set of subsequences of $\{s_k\}$ of measure one corresponding to those $t = \cdot a_1 a_2 \ldots$ in the binary expansion such that $\{a_k\}$ is summable to $\frac{1}{2}$ by $(C, 1)$, then denoting the subsequence of $\{s_k\}$ corresponding to t by $\{s_k(t)\}$, we have

$$\liminf_{r\to\infty} \sum a_{2r,k}\, s_k(t) \geqslant \tfrac{1}{2},$$

$$\lim_{r\to\infty} \sum a_{2r+1,k}\, s_k(t) = 0$$

for almost all t. Hence, although A is stronger than $(C, 1)$ it does not sum almost all subsequences of every bounded A-summable sequence. However A will sum almost all subsequences of a $(C, 1)$-summable sequence. In fact, Buck and Pollard (**3**) pointed out that the set of summable sequences fall in two classes: the one for which the set of summable subsequences is of measure one, the other of measure zero [see also (**14**)].

We now prove the theorem:

THEOREM 3. *If (R, p_n) is a regular Riesz means which satisfies the additional conditions*

$$p_{n+1} \geqslant p_n > 0 \textit{ for all } n = 0, 1, 2, \ldots, \tag{1}$$

$$\sum \left(\frac{p_n}{P_n}\right)^2 \textit{ converges}, \tag{2}$$

$$\frac{p_n P_{k_n}}{p_{k_n} P_n} < L \quad \textit{for all } n \geqslant 0 \textit{ and } k_n = 0, 1, 2, \ldots, 3n, \tag{3}$$

then almost all subsequences of a bounded (R, p_n) summable sequence are (R, p_n) summable to the same sum.

Proof. Consider $$\tau_n(t) = \frac{\sum_0^n p_k(1+R_k(t))s_k}{\sum_0^n p_k(1+R_k(t))}.$$

This is of the form $$\frac{p_{k_0}s_{k_0}+\ldots+p_{k_n}s_{k_n}}{p_{k_0}+\ldots+p_{k_n}}$$

and defines a Riesz means (R, p_{k_n}) of the sequence $\{s_{k_n}\}$. We shall first show that, for almost all subsequences $\{s_{k_n}\}$ of a bounded sequence, $\{s_{k_n}\}$ is summable by the corresponding (R, p_{k_n}). Since

$$\sum \left(\frac{s_n p_n}{P_n}\right)^2 < \infty,$$

then $$\sum \frac{s_n p_n R_n(t)}{P_n}$$

is convergent for almost all t by Lemma 2. By Lemma 3, we then have

$$\sum_0^n s_k p_k R_k(t) = o(P_n), \qquad \sum_0^n p_k R_k(t) = o(P_n)$$

for almost all t. However,

$$\tau_n(t) = \frac{\sum_0^n p_k s_k + \sum_0^n p_k s_k R_k(t)}{P_n} \cdot \frac{P_n}{P_n + \sum_0^n p_k R_k(t)}, \tag{4}$$

and it follows that $$\lim_{n\to\infty} \tau_n(t) = s$$

for almost all t.

The theorem will be established when it is shown that, for almost all t, the methods (R, p_{k_n}) are contained in the method (R, p_k) since it will then follow automatically that all bounded sequences summed by (R, p_{k_n}) will be summed to the (R, p_k) sum [see (**1**) or (**15**)].

Hence, the theorem will follow if, for almost all t [see Hardy (**6**)],

$$\sum_0^{s-1} \left|\frac{p_n}{p_{k_n}} - \frac{p_{n+1}}{p_{k_{n+1}}}\right| \frac{P_{k_n}^+}{P_s} < K \quad \text{and} \quad \frac{p_s P_{k_s}^+}{p_{k_s} P_s} < M, \tag{5}$$

where K and M are constants and

$$P_{k_s}^+ = \sum_0^s p_{k_n}.$$

It easily follows that $P_{k_s}^+/P_{k_s} < M'$. Moreover, from Theorem 2 we see that, for a set of measure one, $k_s \leqslant 3s$ for every $s > N$, so that, from (3),

$$\frac{p_s P_{k_s}^+}{p_{k_s} P_s} = \frac{p_s P_{k_s}}{p_{k_s} P_s} \frac{P_{k_s}^+}{P_{k_s}} \leqslant M'L = M \tag{6}$$

for all $k_s \leqslant 3s$. This means that the second condition for inclusion in (5) is satisfied almost everywhere.

For the first condition

$$U = \sum_0^{s-1} \left|\frac{p_n}{p_{k_n}} - \frac{p_{n+1}}{p_{k_{n+1}}}\right| \frac{P^+_{k_n}}{P_s}$$

$$\leqslant \sum_0^{s-1} \frac{P^+_{k_n}}{P_s} \left| p_n \frac{(p_{k_{n+1}} - p_{k_n})}{p_{k_n} p_{k_{n+1}}} \right| + \sum_0^{s-1} \frac{P^+_{k_n}}{p_{k_n}} \left| \frac{p_n - p_{n+1}}{P_s} \right|$$

$$= U_1 + U_2, \quad \text{say.}$$

Using (6) we have, for almost all t,

$$U_2 \leqslant M \sum_0^{s-1} \frac{P_n}{p_n} \left| \frac{p_n - p_{n+1}}{P_s} \right|$$

and this is bounded, by (3) and Lemma 4. Also

$$U_1 = \sum_0^{s-1} \frac{p_n P^+_{k_n}}{P_s} \left| \frac{1}{p_{k_n}} - \frac{1}{p_{k_{n+1}}} \right|$$

and, since $p_{k_{n+1}} \geqslant p_{k_n} > 0$ for all $n = 0, 1, 2,\ldots$,

$$U_1 \leqslant \sum_0^{s-1} \frac{1}{p_{k_n} P_s} (p_n P^+_{k_n} - p_{n-1} P^+_{k_{n-1}}) + \frac{p_{s-1} P^+_{k_{s-1}}}{p_{k_s} P_s}$$

with the convention that $p_{-1} P^+_{-1} = 0$. Again, using (6) and remembering that $p_{s-1} P^+_{k_{s-1}} \leqslant p_s P^+_{k_s}$, we have

$$U_1 \leqslant \sum_0^{s-1} \frac{p_{n-1}}{p_{k_n} P_s} [P^+_{k_n} - P^+_{k_{n-1}}] + \sum_0^{s-1} \frac{p_n - p_{n-1}}{p_{k_n}} \frac{P^+_{k_n}}{P_s} + M.$$

Here we have

$$\sum_0^{s-1} \frac{p_{n-1}}{p_{k_n} P_s} [P^+_{k_n} - P^+_{k_{n-1}}] = \sum_0^{s-1} \frac{p_{n-1}}{p_{k_n} P_s} p_{k_n} = \frac{1}{P_s} \sum_0^{s-1} p_{n-1} \leqslant 1,$$

and from (6), for almost all t,

$$\sum_0^{s-1} \frac{P^+_{k_n}}{p_{k_n}} \frac{p_n - p_{n-1}}{P_s} \leqslant M \sum_0^{s-1} \frac{P_{n-1}}{p_{n-1}} \frac{p_n - p_{n-1}}{P_s}.$$

The second sum is bounded, by Lemma 4. Hence U_1 is bounded and the first part of (5) is satisfied for almost all t. This completes the proof of the theorem.

Lorentz (**11**) has introduced the concept of 'strongly regular methods'. A matrix $A = (a_{mn})$ is *strongly regular* if and only if

$$\lim_{m\to\infty} \sum |a_{m,n} - a_{m,n+1}| = 0.$$

Associated with strongly regular methods are summability functions; a positive monotonically increasing function $\Omega(n) \nearrow \infty$ such that

$$\lim_{m\to\infty} \sum \Omega(n) |a_{m,n} - a_{m,n+1}| = 0$$

is called a *summability function* of the method $A = (a_{mn})$.

Lorentz (**13**) proved the lemma:

LEMMA 5. *If $f(n)$ is a positive monotonically increasing function such that every $\Omega(n) = o(f(n))$ is a summability function of A, then*

$$\sum f(n) |a_{m,n} - a_{m,n+1}| \leqslant M.$$

If $p_k \geqslant p_{k-1} > 0$ for $k = 0, 1, 2,...,$ the regular Riesz method (R, p_k) is strongly regular if and only if

$$\lim_{m\to\infty} \frac{p_n}{P_n} = 0.$$

For such methods, Lorentz (**12**) proved the theorem:

THEOREM 4. *If $p_n \geqslant p_{n-1} > 0$ for all $n = 1, 2, 3,...$ and p_n/P_n decreases to zero then any function of the form $o(P_n/p_n)$ is a summability function for (R, p_n) means.*

It is clear from Lemma 5 and Theorem 4 that, if $p_n \geqslant p_{n-1} > 0$ for all n and p_n/P_n decreases to 0, then

$$\sum_{0}^{n-1} \frac{P_k}{p_k} \left| \frac{p_k - p_{k+1}}{P_n} \right| \leqslant M$$

for all $n = 0, 1, 2,....$ Theorem 4 also follows immediately since p_n/P_n decreases to 0.

Let us assume that

$$\frac{P_n p_{k_n}}{p_n P_{k_n}} < L' \quad (\tfrac{1}{3}n \leqslant k_n \leqslant 3n). \tag{3'}$$

Then we prove a generalization of another theorem due to Buck and Pollard (**3**).

THEOREM 5. *If (R, p_k) is a regular method satisfying the conditions of Theorem 4 and in addition p_n/P_n decreases to 0 and (3′) holds, then a bounded sequence that is not (R, p_k)-summable has almost no subsequences that are (R, p_k)-summable.*

Proof. We have
$$\frac{P_n p_{k_n}}{p_n P^+_{k_n}} = \frac{P_n p_{k_n}}{p_n P_{k_n}} \frac{P_{k_n}}{P^+_{k_n}}.$$

However,
$$\frac{P_{k_n}}{P^+_{k_n}} = \frac{\sum_0^n p_r}{\sum_0^n p_{k_r}} = \frac{2\sum_0^{k_n} p_r}{\sum_0^{k_n} p_r + \sum_0^{k_n} p_r R_r(t)}$$

and, from considerations similar to those in (4), $P_{k_r}/P^+_{k_n}$ is bounded for almost all subsequences. Finally, it is clear that $\frac{1}{3}n \leqslant k_n \leqslant 3n$ for almost all subsequences. It then follows that
$$\frac{P_n p_{k_n}}{p_n P^+_{k_n}} < M$$

for almost all subsequences. We can now prove that
$$\sum_0^{s-1} \left|\frac{p_{k_n}}{p_n} - \frac{p_{k_{n+1}}}{p_{n+1}}\right| \frac{P_n}{P^+_{k_s}} < K$$

in much the same way as we proved the first part of (5). Here, however, we must replace Lemma 4 by Theorem 4 since the conditions of the lemma are not satisfied by (3′). We now have that (R, p_{k_n}) is stronger than (R, p_k) for almost all t. Hence, if $\{s_k\}$ is a bounded sequence that is not (R, p_k)-summable,
$$\sum_0^n s_k p_k R_k(t) = o(P_n)$$

for almost all t; and, from (4) again, for almost all t,
$$\frac{p_{k_0} s_{k_0} + \dots + p_{k_n} s_{k_n}}{p_{k_0} + \dots + p_{k_n}}$$

does not converge. But (R, p_{k_n}) is stronger than (R, p_k), and hence, for almost all t, the subsequences are not (R, p_k)-summable. This completes the proof of the theorem.

The conditions of Theorem 3 can easily be seen to be satisfied by the Cesàro means, i.e. the (R, p_k)-means where $p_k = 1$ for all k. The conditions are also satisfied by the (R, p_k)-means defined by $P_k = \exp(\log^2 k)$ for all $k = 1, 2, 3, \dots$. On the other hand (**12**) this Riesz means has all summability functions $o(n/\log n)$ and does not sum all bounded $(C, 1)$-summable sequences.

It has been shown by Lorentz (**13**) that, if $A = (a_{mn})$ is a strongly regular matrix and $f(n)$ a positive monotonically increasing function such that $\sum f(n)|a_{mn} - a_{m,n+1}| \leqslant M$ for all m, then A is stronger than

the (R, p_k) means defined by

$$P_k = e^{\mu_k}, \quad \text{where} \quad \mu_k = \frac{1}{f(1)} + \dots + \frac{1}{f(k-1)}.$$

The referee of this paper has pointed out to the authors that Lorentz's theorem [(**13**) 131, Theorem 2] includes some conditions not stated but implied by previous work. We remark that $f(n) \nearrow \infty$ implies that p_n/P_n decreases to zero, and that these implied conditions for A to contain (R, p_n) are fulfilled. For further details see the references quoted in (**13**).

For such a method

$$\frac{p_k}{P_k} = 1 - \exp\left(-\frac{1}{f(k)}\right).$$

Since $f(k)$ is positive increasing, we have $f(k) \geqslant \text{constant} > 0$, and hence the expression on the right lies between two positive constant multiples of $1/f(k)$. Hence $\sum (p_k/P_k)^2$ will converge if and only if (i) $\sum f^{-2}(k)$ converges. It is clear also that condition (3) of Theorem 3 is satisfied if (ii) $f(k_s)/f(s) < L$ for all $k_s \leqslant 3s$. Hence, if $f(n)$ satisfies (i) and (ii), $A = (a_{mn})$ is stronger than a Riesz means (R, p_k) satisfying the conditions of the theorem. The method A will sum all bounded (R, p_k) summable sequences (and to the same sum (**1**), (**15**)), and hence sums almost all subsequences of these bounded sequences to the same sum as the original sequences. One can speak of the 'Buck–Pollard subset of bounded B-summable sequences', where $B = (b_{mn})$ is any regular matrix, meaning those bounded B-summable sequences for which almost all subsequences are summable to the same value. The Buck–Pollard subset is non-empty for strongly regular methods [see (**14**)]. For the Riesz means of our theorem and $(C, 1)$, the Buck–Pollard subset is the set of bounded summable sequences. For methods such as the $A = (a_{mn})$ described immediately above, the Buck–Pollard subset does not always coincide with the set of summable sequences. The category of the set of summable subsequences is discussed in (**9**) and (**10**).

In Theorem 3 there is no need to assume $\{s_n\}$ bounded: the argument applies under the weaker assumption that

$$\sum \left(\frac{s_n p_n}{P_n}\right)^2 < \infty.$$

We now wish to treat with some properties of a matrix introduced by Garreau (**5**). The matrix $G = (g_{mn})$ is described as follows: let the

first two rows consist of 1, 0 in both orders, followed by zeros; the next ${}_4C_2$ (= 6) rows of 0, 0, $\frac{1}{2}$, $\frac{1}{2}$ in every possible order, followed by zeros; the next ${}_6C_3$ (= 20) rows of 0, 0, 0, $\frac{1}{3}$, $\frac{1}{3}$, $\frac{1}{3}$, in every possible order followed by zeros and so on, so that there are ${}_{2v}C_v$ rows consisting of v zeros and v elements equal to v^{-1} arranged in every possible order followed by zeros. This is a regular matrix and, moreover,

$$\lim_{m\to\infty} \max |a_{mn}| = 0.$$

Hill (**8**) made an extensive investigation of the regular matrices which sum almost all sequences of 0's and 1's, i.e. that have the Borel property. He stated that a necessary condition for a matrix to have the Borel property is that

$$\lim_{m\to\infty} \sum a_{mn}^2 = 0.$$

Garreau used the matrix G to show that this condition is not sufficient, i.e. G satisfies this condition but does not have the Borel property. We shall now show that G is an example of a matrix that sums a bounded sequence if and only if it sums almost all of its subsequences. Hence we have a matrix which has the Buck–Pollard property but not the Borel property. A matrix stronger than a matrix with the Borel property must also have the Borel property. It is clear that the matrix we have constructed in the discussion following Theorem 2, being stronger than the Cesàro matrix, has the Borel property. Hence, this is an example of a matrix having the Borel property but not the Buck–Pollard property. From these two examples it is clear that the Borel property and the Buck–Pollard property are independent.

If $\nu(n)$ be the number of s_k ($k \leqslant n$) such that $s_k \in E$, where E is some set of real numbers, then $\nu(n)$ will be called a 'counting function' relative to E for the sequence $\{s_n\}$. If E is the set of non-zero real numbers, we shall call $\nu(n)$ the 'counting function' for $\{s_n\}$. If the counting function $\nu(n) = o(n)$ and $\{s_n\}$ is bounded, then it is evident that $\{s_n\}$ is G summable to zero. Hence any sequence that is the sum of a sequence convergent to zero and one whose counting function is $o(n)$ is G summable to zero. We now prove the theorem:

THEOREM 6. *A bounded sequence $\{s_n\}$ is G summable to zero if and only if $s_n = x_n + y_n$ where $\{x_n\}$ converges to zero and $\{y_n\}$ has a counting function $\nu(n) = o(n)$.*

Our previous remarks show the sufficiency of the conditions. On the other hand, suppose that $\{s_n\}$ is G-summable and bounded. Let $\nu(\epsilon, n)$ be defined as the counting function of E (E: $|x| > \epsilon$). We first show

that, for any $\epsilon > 0$, $\nu(\epsilon, n) = o(n)$. For, if not, there exists a sequence $\{2n_v\}$ and an ϵ such that the counting function relative to the set $(E: x > \epsilon)$ has $\nu(2, n_v) > \alpha n_v$ $(v = 1, 2,...)$ for some $\alpha > 0$ (or else this will be true relative to the set $(E: x < -\epsilon)$). If $\alpha > 1$, we choose the row corresponding to that one of the ${}_{2n_v}C_{n_v}$ combinations for which each non-zero element of the row corresponds to an $s_r > \epsilon$. For such a row, m_μ say,

$$\sum g_{m_\mu n} s_n > \epsilon.$$

If $\alpha < 1$, there are three cases to consider.

(i) The row may be so chosen that the remaining $(1-\alpha)n_v$ terms of the row where $g_{m_\mu n} \neq 0$ correspond to values where $s_n > 0$ $(n < 2n_v)$. In this case, we again have $\sum g_{m_\mu n} s_n > \alpha\epsilon$.

(ii) If a choice as in (i) is not possible owing to an insufficient number of $s_n \geqslant 0$ $(n \leqslant 2n_v)$, we choose a row so that the remaining $(1-\alpha)n_v$ terms of the row where $g_{m_\mu n} \neq 0$ correspond to s_n such that the sum of the negative terms does not exceed $\frac{1}{2}\alpha\epsilon$ in absolute value. If such a choice is possible,

$$\sum g_{m_\mu n} s_n \geqslant \tfrac{1}{2}\alpha\epsilon.$$

(iii) If the choice indicated in (i) or (ii) is not possible, then over half the terms of the sequence for which $n < 2n_v$ are negative. Moreover, the sum of $(1-\alpha)n_v$ $(0 < \alpha < 1)$ of them is less than $-\frac{1}{2}\alpha\epsilon$, and so it is clear that a row can be found so that the $g_{m_\mu n} \neq 0$ all correspond to $s_n < 0$ and so that $\sum g_{m_\mu n} s_n < -\frac{1}{2}\alpha\epsilon$. Hence, in any case we can select a sequence $\{m_\mu\}$ such that

$$\left|\sum g_{m_\mu n} s_n\right| > \tfrac{1}{2}\alpha\epsilon$$

for all $\mu = 1, 2, 3,...$. Therefore, we have $\nu(\epsilon, n) = o(n)$ for any $\epsilon > 0$ since an assumption to the contrary leads to a contradiction.

It now follows that we can choose an increasing sequence of integers $\{n_k\}$ such that

$$\nu(2^{-k}, n) \leqslant 2^{-k}n \quad (n > n_k).$$

Let $x_n = s_n$ if $|s_n| \leqslant 2^{-k}$, $x_n = 0$ if $|s_n| \geqslant 2^{-k}$, where $n_k \leqslant n \leqslant n_{k+1}$. Then it is clear that $\{x_n\}$ converges and that the counting function for $\{s_n - x_n\} = \{y_n\}$ is $o(n)$. This completes the proof of the theorem.

We now prove the theorem:

THEOREM 7. *A bounded sequence is G-summable if and only if almost all of its subsequences are G-summable to the same sum.*

This theorem will clearly be proved if we can prove that a bounded sequence is summable by G to zero if and only if almost all its subsequences are G-summable to zero.

Let us suppose first of all that the counting function of $\{s_n\}$ satisfies the condition $\nu(n) = o(n)$. Define

$$t_n = \begin{cases} 1 & (s_n \neq 0), \\ 0 & (s_n = 0). \end{cases}$$

If the counting function of $\{s_n\}$ is $o(n)$, then $\{t_n\}$ is summable $(C, 1)$ to 0. Hence, so are almost all subsequences. But any subsequence of $\{t_n\}$ consists of 0's and 1's; and for a sequence of 0's and 1's to be summable $(C, 1)$ to 0, it is necessary that the counting function be $o(n)$. This means that almost all subsequences of $\{t_n\}$ have a counting function $o(n)$, and the same is true for $\{s_n\}$. This means that almost all of the subsequences of $\{s_n\}$ are G-summable to zero. Since every subsequence of a convergent sequence is convergent to the same sum, it is clear by splitting a G-summable sequence as in Theorem 6 that almost all its subsequences are G-summable. On the other hand, if almost all the subsequences of $\{s_n\}$ are G-summable to zero, almost all the subsequences of $\{|s_n|\}$ are G-summable to zero. The Cesàro matrix is a submatrix of G, so that almost all of the subsequences of $\{|s_n|\}$ must also be Cesàro-summable to zero. This means again that $\{|s_n|\}$ is Cesàro-summable to zero, and it is clear that it must be possible to split $\{|s_n|\}$ as in Theorem 6. The same will be true of $\{s_n\}$, so that $\{s_n\}$ must be G-summable to zero. This completes the proof of our theorem.

The authors are grateful to the referee for his painstaking care, and especially for widening the conditions of Theorem 3.

REFERENCES

1. A. Brudno, 'Summation of bounded sequences by matrices', *Mat. Sbornik* 16 (1945) 191–247 (in Russian).
2. R. C. Buck, 'A note on subsequences', *Bull. American Math. Soc.* 49 (1943) 898–9.
3. —— and H. Pollard, 'Convergence and summability properties of subsequences', ibid. 924–31.
4. R. G. Cooke, *Infinite matrices and sequence spaces* (London 1950) 207–13.
5. G. A. Garreau, 'A note on the summation of sequences of 0's and 1's', *Annals of Math.* (2) 54 (1951) 183–5.
6. G. H. Hardy, *Divergent series* (Oxford, 1949) 58.
7. J. D. Hill, 'Remarks on the Borel property', *Pacific J. of Math.* 4 (1954) 227–42.
8. —— 'Summability of sequences of 0's and 1's', *Annals of Math.* 46 (1945) 556–62.
9. F. R. Keogh and G. M. Petersen, 'A universal Tauberian theorem', *J. London Math. Soc.* 33 (1958) 121–3.
10. —— —— 'A generalized Tauberian theorem,' *Canadian J. of Math.* 10 (1958) 111–14.

11. G. G. Lorentz, 'A contribution to the theory of divergent sequences', *Acta Math.* 80 (1948) 167–90.

12. —— 'Direct theorems on methods of summability' (I), (II), *Canadian J. of Math.* 1 (1949) 304–19; 3 (1951) 236–56.

13. —— 'Borel and Banach properties of methods of summation', *Duke Math. J.* 22 (1955) 129–41.

14. G. M. Petersen, 'Almost convergence and the Buck–Pollard property', *Proc. American Math. Soc.* 11 (1960) 469–77.

15. —— 'Summability methods and bounded sequences', *J. London Math. Soc.* 31 (1956) 324–6.

FUNCTION TO SEQUENCE MAPPING

By S. N. PRASAD (*London*)

[Received 20 February 1960]

1. Introduction

THE object of this paper is to consider transformations of functions to sequences, sets of functions to sets of sequences, and function spaces into sequence spaces. Transformations of sequence to sequence are effected by a matrix $(a_{n,k})$ [see (**1**) chapter 4, (**2**) chapter 6, and (**3**)], sequence to function transformations are effected by $(a_k(\omega))$, where ω is a positive continuous variable [see (**1**) chapter 4], and series to series transformations are carried out by matrices of a different type [see (**4**) and (**5**)]. For transformations from function to sequence we shall employ a sequence of functions $\{F_n(x)\}$; and $F_n(x)$ will be referred to as a *transformer*.

All the variables considered are real, and all the functions and transformers considered are real and measurable in the Lebesgue sense in $[0, a]$ for all $a > 0$. The integration is in the Lebesgue sense throughout.

The transformer $F_n(x)$ is said to *apply absolutely* to $g(x)$ and to *transform* $g(x)$ if

$$u_n(g) = \int\limits_0^\infty F_n(x)g(x)\,dx \tag{1.1}$$

is absolutely convergent for each n. Then $\{u_n(g)\}$ is called the *sequence transform of the function* $g(x)$. If $F_n(x)$ applies absolutely to every $g(x)$ in a function space $\alpha(f)$, it is said to *apply absolutely to the function space* $\alpha(f)$, and the sequence transforms $\{u_n(g)\}$ of the functions $g(x)$ in $\alpha(f)$ form a sequence space; and, if this sequence space is β, we say that the *function space* $\alpha(f)$ has been *transformed into the sequence space* β.

2. Definitions and notations

Definitions of the sequence spaces used in this paper are given in [(**1**) 273–4], and those of the function spaces are to be found in [(**6**) § 3].

Sequence spaces will be denoted by α, β,..., and function spaces by $\alpha(f)$, $\beta(f)$,.... $\{f_n(x)\}$ will mean a sequence of functions $f_1(x)$, $f_2(x)$, $f_3(x)$,..., defined for almost all $x \geqslant 0$.

The sequence of functions $\{f_n(x)\}$ is said to be *coordinate convergent* (c-cgt) if $f_n(x)$ tends to a finite limit as $n \to \infty$ for almost all $x \geqslant 0$.

The sequence of functions $\{f_n(x)\}$ is said to be *uniformly coordinate convergent* (c-cgt(u)) if $f_n(x)$ converges uniformly to a finite limit as $n \to \infty$ for almost all $x \geqslant 0$, i.e. if to every $\epsilon > 0$ there corresponds $N(\epsilon)$, *independent of* x, such that, for almost all $x \geqslant 0$,

$$|f_n(x) - f_{n'}(x)| \leqslant \epsilon \tag{2.1}$$

for all n, $n' \geqslant N(\epsilon)$.

For the definition of the dual space $\alpha^*(f)$ of a function space $\alpha(f)$, we refer to (**6**).

If $\alpha^*(f) \geqslant \beta(f)$, then $f_n(x)$ in $\alpha(f)$ is said to be *projective convergent* (p-cgt) *relative to* $\beta(f)$, or $\alpha(f)\beta(f)$-cgt, if to every $\epsilon > 0$ and every $g(x)$ in $\beta(f)$ there corresponds $N(\epsilon, g)$ such that

$$\left| \int_0^\infty g(x)\{f_n(x) - f_{n'}(x)\}\, dx \right| \leqslant \epsilon$$

for all n, $n' \geqslant N(\epsilon, g)$.

When $\beta(f) = \alpha^*(f)$, $f_n(x)$ is said to be 'p-cgt in $\alpha(f)$', or '$\alpha(f)$-cgt'.

Taking $f_n(x) = \cos nx$ and $g(x)$ to be any function in $\sigma_1(f)$, we see, by the Riemann–Lebesgue theorem, that $f_n(x)$ is $\sigma_\infty(f)$-cgt. But $\cos nx$ does not tend to a limit as $n \to \infty$. Hence $\sigma_\infty(f)$-convergence does not necessarily imply c-convergence; i.e. in general, $\alpha(f)\beta(f)$-convergence does not necessarily imply coordinate convergence.

A theorem will now be proved which will be required in the subsequent theory.

$\{f_n(x)\} \in \mathscr{M}$ will mean that $\{f_n(x)\}$ is *monotonic increasing* or *monotonic decreasing* for almost all $x \geqslant 0$.

(2, I). *If* $\beta(f) \geqslant \phi(f)$, $\{f_n(x)\} \in \mathscr{M}$, *and* $f_n(x)$ *is* $\alpha(f)\beta(f)$-cgt, *then* $\{f_n(x)\}$ *is* c-cgt(u).

The proof is similar to that of [(**6**) (**4**, III)].

$[\alpha(f) \to \beta]$ will denote the space of all $F_n(x)$ which transform $\alpha(f)$ into β; these $F_n(x)$ form a function space.

$F_n(x)$ belonging to $\sigma_1(f)$ is said to be *convergent in mean* if to every $\epsilon > 0$ there corresponds a number $N(\epsilon)$ such that

$$\int_0^\infty |F_n(x) - F_{n'}(x)|\, dx \leqslant \epsilon \tag{2.2}$$

for all n, $n' \geqslant N(\epsilon)$.

$F_n(x)$ in $\sigma_1(f)$ is said to be *convergent in mean to zero* if to every $\epsilon > 0$ there corresponds a number $N(\epsilon)$ such that, for all $n \geqslant N(\epsilon)$,

$$\int_0^\infty |F_n(x)|\,dx \leqslant \epsilon.$$

A set X in $\sigma_1(f)$ will be said to be *integral-bounded* (int-bd) if

$$\int_0^\infty |g(x)|\,dx \leqslant M$$

for every $g(x)$ in X, where M is a constant independent of g.

A set X in $\sigma_\infty(f)$ will be said to be *collectively bounded* when $|f(x)| \leqslant M$ for almost all $x \geqslant 0$ and for every f in X.

For the definition of a *collectively convergent* set in Γ we refer to [(1) 282].

A set X in Z is *collectively convergent* when to every $\epsilon > 0$ corresponds a number $N(\epsilon)$ such that, for every $\{x_k\}$ in X, $|x_k| \leqslant \epsilon$ for all $k \geqslant N(\epsilon)$.

We refer to [(1) 278] for the definition of a *normal* sequence space.

3. Some theorems on transformations

(3, I). *If $F_n(x) \geqslant 0$ for almost all $x \geqslant 0$ and for every n, then $F_n(x)$ transforms $\alpha(f) \leqslant \sigma_\infty(f)$ into a normal space β if and only if*

$$\text{(i) } F_n(x) \in \sigma_1(f) \quad \textit{and} \quad \text{(ii) } \left\{\int_0^\infty F_n(x)\,dx\right\} \in \beta,$$

where $\alpha(f)$ contains the constant functions.

Since $\alpha(f) \leqslant \sigma_\infty(f)$ and contains the constant functions, $\alpha^*(f) = \sigma_1(f)$, by [(6) (3, I) and (3, II)].

Let $g(x)$ be any function in $\alpha(f)$ and let $u_n(g)$ be given by (1.1).

Suppose that $F_n(x) \in \sigma_1(f)$ and

$$\left\{\int_0^\infty F_n(x)\,dx\right\} \in \beta.$$

Since $F_n(x) \in \sigma_1(f) = \alpha^*(f)$, $F_n(x)$ applies absolutely to $\alpha(f)$, and the integral in (1.1) exists for every n.

Now $$|u_n(g)| = \left|\int_0^\infty F_n(x)g(x)\,dx\right| \leqslant M\int_0^\infty F_n(x)\,dx, \tag{3.1}$$

since $\alpha(f) \leqslant \sigma_\infty(f)$, and $F_n(x) \geqslant 0$ for almost all $x \geqslant 0$ and for every n.

From (3.1), we see that $\{u_n(g)\} \in \beta$ since

$$\left\{\int_0^\infty F_n(x)\,dx\right\} \in \beta$$

and β is normal. Thus the conditions are sufficient.

Now suppose that $\{u_n(g)\}$ exists for every g in $\alpha(f)$ and belongs to β. Then $\{u_n(g)\}$ is the sequence transform of g by $F_n(x)$, where g is any function in $\alpha(f)$. So $F_n(x) \in [\alpha(f) \to \beta]$, and consequently

$$F_n(x) \in \alpha^*(f) = \sigma_1(f).$$

Also
$$\{u_n(g)\} = \left\{ \int_0^\infty F_n(x) g(x)\,dx \right\} \in \beta$$

for every g in $\alpha(f)$, and, taking $g = 1$ everywhere in $[0, \infty)$, we see that

$$\left\{ \int_0^\infty F_n(x)\,dx \right\} \in \beta.$$

Hence the conditions are necessary.

Examples. The spaces ϕ, σ_r $(r \geqslant 1)$, σ_∞, σ, Z, O_1, O_2, $\bar{O}_1$, $\bar{O}_2$, E_r, F_r, and δ are normal [(**1**) 278], so that, in (3, I), β could be any one of these spaces.

(3, II). *If* $\phi(f) \leqslant \beta(f) \leqslant \sigma_1(f)$, *and* $\{F_n(x)\} \in \mathcal{M}$, *then* $F_n(x)$ *transforms* $\beta(f)$ *into* Γ *if and only if* (i) $F_n(x) \in \beta^*(f)$ *and* (ii) $\{F_n(x)\}$ *is* c-cgt(u).

Let $g(x) \in \beta(f)$ and let $u_n(g)$ be given by (1.1). If $F_n(x) \in \beta^*(f)$, F_n applies absolutely to $\beta(f)$ and $\{u_n(g)\}$ exists for every g in $\beta(f)$.

If also $\{F_n(x)\}$ is c-cgt(u), then, by (2.1), given any $\epsilon > 0$,

$$|F_n(x) - F_{n'}(x)| \leqslant \epsilon \tag{3.2}$$

for all n, $n' \geqslant N(\epsilon)$ and for almost all $x \geqslant 0$; N is independent of x. Then, by (1.1) and (3.2),

$$|u_n(g) - u_{n'}(g)| = \left| \int_0^\infty \{F_n(x) - F_{n'}(x)\} g(x)\,dx \right|$$

$$\leqslant \epsilon \int_0^\infty |g(x)|\,dx \leqslant \epsilon k(g)$$

for all n, $n' \geqslant N(\epsilon)$ since $\beta(f) \leqslant \sigma_1(f)$. Hence $\{u_n(g)\} \in \Gamma$ for every g in $\beta(f)$. Thus the conditions are sufficient.

Conversely, if $\{u_n(g)\}$ exists for every g in $\beta(f)$ and belongs to Γ, then $F_n(x) \in [\beta(f) \to \Gamma]$, and so $F_n(x) \in \beta^*(f)$; also, with any $\epsilon > 0$, there exists $N(\epsilon, g)$ such that, for all n, $n' \geqslant N(\epsilon, g)$ and for every g in $\beta(f)$,

$$\left| \int_0^\infty \{F_n(x) - F_{n'}(x)\} g(x)\,dx \right| \leqslant \epsilon.$$

Hence $F_n(x)$ is $\beta^*(f)\beta(f)$-cgt. But, by hypothesis, $\beta(f) \geqslant \phi(f)$ and

$\{F_n(x)\} \in \mathscr{M}$; therefore, by (2, I), $\{F_n(x)\}$ is c-cgt(u). Hence the conditions are necessary.

(3, III). *$F_n(x)$ belonging to $\sigma_1(f)$ transforms collectively bounded sets in $\sigma_\infty(f)$ into collectively convergent sets in Γ if and only if $F_n(x)$ is convergent in mean.*

Let $g(x) \in X$, a collectively bounded set in $\sigma_\infty(f)$; then, by definition, $|g(x)| \leqslant M$ for almost all $x \geqslant 0$ and for every g in X.

Let $u_n(g)$ be given by (1.1), and suppose that $F_n(x)$ in $\sigma_1(f)$ is convergent in mean; then, by (2.2), for any $\epsilon > 0$, there exists $N(\epsilon)$ such that

$$\int\limits_0^\infty |F_n(x)-F_{n'}(x)|\, dx \leqslant \epsilon$$

for all $n, n' \geqslant N(\epsilon)$, and therefore

$$|u_n(g)-u_{n'}(g)| = \left|\int\limits_0^\infty \{F_n(x)-F_{n'}(x)\}g(x)\, dx\right|$$

$$\leqslant M \int\limits_0^\infty |F_n(x)-F_{n'}(x)|\, dx \leqslant M\epsilon$$

for all n, $n' \geqslant N(\epsilon)$ and for every g in X.

Hence the sequences $\{u_n(g)\}$ form a collectively convergent set in Γ.

Conversely, if the sequences $\{u_n(g)\}$ form a collectively convergent set in Γ, then

$$\left|\int\limits_0^\infty \{F_n(x)-F_{n'}(x)\}g(x)\, dx\right| \leqslant \epsilon \tag{3.3}$$

for all n, $n' \geqslant N(\epsilon, X)$, for every $\epsilon > 0$, and for every g in X, where X is any collectively bounded set in $\sigma_\infty(f)$.

If $F_n(x)$ is not convergent in mean, then there is some $\epsilon > 0$ for which

$$\int\limits_0^\infty |F_n(x)-F_{n'}(x)|\, dx > \epsilon \tag{3.4}$$

for an infinity of pairs n, n'.

Let us take the collectively bounded set X' consisting of the functions $g(x)$ such that $|g(x)| = 1$ for all $x \geqslant 0$.

Corresponding to an ϵ for which (3.4) holds and to the set X', there exists an $N(\epsilon, X')$ for which (3.3) is true.

Pick out any pair n, n' for which both (3.3) and (3.4) hold, and for this fixed pair n, n' take $g(x) = 1$ or -1 according as $F_n(x) \geqslant$ or $< F_{n'}(x)$,

so that $g(x)$ is in X', and, by (3.3),

$$\int_0^\infty |F_n(x)-F_{n'}(x)|\,dx \leqslant \epsilon \tag{3.5}$$

for this fixed pair n, n'.

But (3.4) contradicts (3.5); hence (3.4) is impossible, and so $F_n(x)$ must be convergent in mean. We have the corollary:

COROLLARY. *$F_n(x)$ belonging to $\sigma_1(f)$ transforms collectively bounded sets in $\sigma_\infty(f)$ into collectively convergent sets in Z if and only if $F_n(x)$ is convergent in mean to* 0.

(3, IV). *$F_n(x)$ belonging to $\sigma_\infty(f)$ transforms* int-bd *sets in $\sigma_1(f)$ into collectively convergent sets in* Γ *if and only if* $\{F_n(x)\}$ *is* c-cgt(u).

Let X be an int-bd set in $\sigma_1(f)$; then, by definition,

$$\int_0^\infty |g(x)|\,dx \leqslant M \tag{3.6}$$

for every g in X.

Let $\{F_n(x)\}$ be c-cgt(u); then, by (2.1), with any $\epsilon > 0$, there exists $N(\epsilon)$ such that, for almost all $x \geqslant 0$,

$$|F_n(x)-F_{n'}(x)| \leqslant \epsilon \tag{3.7}$$

for all $n, n' \geqslant N(\epsilon)$.

Let $g(x) \in X$, and $u_n(g)$ be given by (1.1); then, by (3.6) and (3.7),

$$\begin{aligned}|u_n(g)-u_{n'}(g)| &= \left|\int_0^\infty \{F_n(x)-F_{n'}(x)\}g(x)\,dx\right| \\ &\leqslant \int_0^\infty |F_n(x)-F_{n'}(x)|\,|g(x)|\,dx \leqslant \epsilon M\end{aligned}$$

for all $n, n' \geqslant N(\epsilon)$ and for every g in X. Hence the sequences $\{u_n(g)\}$, where $g \in X$, form a collectively convergent set in Γ.

Conversely, if the sequences $\{u_n(g)\}$ form a collectively convergent set in Γ for the functions $g(x) \in X$, where X is any int-bd set in $\sigma_1(f)$, then, for every g in X,

$$\left|\int_0^\infty \{F_n(x)-F_{n'}(x)\}g(x)\,dx\right| \leqslant \epsilon$$

for all $n, n' \geqslant N(\epsilon, X)$ and for every int-bd set in $\sigma_1(f)$. Hence we can prove, by the same method of proof as in (3, III), that $\{F_n(x)\}$ is c-cgt(u).

I wish to thank Dr. R. G. Cooke for some useful suggestions in connexion with the paper.

REFERENCES

1. R. G. Cooke, *Infinite matrices and sequence spaces* (London, 1950).
2. —— *Linear operators* (London, 1953).
3. H. S. Allen, 'Transformations of sequence spaces', *J. London Math. Soc.* 31 (1956) 374–6.
4. P. Vermes, 'Series to series transformations and analytic continuation by matrix methods', *American J. Math.* 71 (1949) 541–62.
5. —— 'Certain classes of series to series transformation matrices', ibid. 72 (1950) 615–20.
6. S. N. Prasad, 'Analogues of some results on sequence spaces for function spaces', *Quart. J. of Math.* (Oxford) (2) 11 (1960) 310–20.

THE SPHERE, CONSIDERED AS AN H-SPACE MOD p

By J. F. ADAMS (*Cambridge*)

[Received 28 May 1960]

I. M. JAMES (4) has suggested a classification of H-spaces into ten classes and has given examples to show that five of these classes are non-empty. It is one of the objects of this paper to give examples showing that all ten classes are non-empty. I begin by summarizing James' classification.

By an *H-space* we shall understand a space X provided with a base-point $e \in X$ and a product map μ: $X \times X \to X$ such that $\mu(x,e) = x$, $\mu(e,x) = x$. We shall say that two H-spaces (X, e, μ), (X', e', μ') are *equivalent* if there is a homotopy-equivalence f: X, $e \to X'$, e' such that the following diagram is homotopy-commutative.

$$\begin{tikzcd} X \times X \arrow[r, "\mu"] \arrow[d, "f \times f"'] & X \arrow[d, "f"] \\ X' \times X' \arrow[r, "\mu'"] & X' \end{tikzcd}$$

Let A denote an adjective applicable to H-spaces, as for example 'associative' or 'commutative'; then we shall say that an H-space (X, e, μ) is *equivalent-A* if there is an equivalent H-space (X', e', μ') which is A.

An H-space may be

(a) not homotopy-associative, or

(b) homotopy-associative but not equivalent-associative, or

(c) equivalent-associative.

Similarly, it may be

(d) not homotopy-commutative, or

(e) homotopy-commutative, but not equivalent-commutative, or

(f) equivalent-commutative.

In this way we obtain nine classes of H-spaces. One of these classes, however, can be divided into two, for, if an H-space is equivalent-associative and also equivalent-commutative, it may be

(g) equivalent-(associative and commutative), or

(h) not so.

In this way H-spaces can be divided into ten classes.

I remark that our definition of 'equivalent-associative' (chosen for its obvious homotopy-invariance) is different from James' definition and not obviously equivalent to it. It is therefore possible that the ten classes defined above differ from James' ten classes. However, the ten examples I shall give serve equally well in either context.

These ten examples are all spheres, mutilated in one fashion or another. To prove that they are not equivalent-associative, or not homotopy-associative, we rely on the following two lemmas:

LEMMA 1. *For each odd integer* $n \geqslant 5$ *there is an infinity of primes* p *such that the exterior algebra* $E(n; Z_p)$ *is not a possible cohomology ring* $H^*(X; Z_p)$ *for an equivalent-associative* H*-space.*

LEMMA 2. *For suitable odd* n *(e.g.* $n = 7$*) the exterior algebra* $E(n; Z_3)$ *is not a possible cohomology ring* $H^*(X; Z_3)$ *for a homotopy-associative* H*-space.*

Proof of Lemma 1. Suppose that X is an equivalent-associative H-space whose cohomology ring is $E(n; Z_p)$. Then there is an equivalent H-space X' which is associative and has the same cohomology ring. By (**2**) it has a classifying space B; and $H^*(B; Z_p)$ is a polynomial ring $P(n+1; Z_p)$. Let x be a generator of this ring; then $x^p \neq 0$: that is, $P_p^f(x) \neq 0$, where $f = \frac{1}{2}(n+1)$. By the Adem relations (**1**), P_p^f can be factorized in terms of operations P_p^g, where $g = p^h$. Therefore at least one such operation P_p^g is non-zero in $P(n+1; Z_p)$. This implies that

$$n+1 \mid 2(p-1)p^h.$$

This is impossible if

$$p > n+1, \qquad p \not\equiv 1 \bmod \tfrac{1}{2}(n+1).$$

If $n \geqslant 5$, we can find an infinity of primes p satisfying these conditions.

The proof of Lemma 2 is closely similar. If X is a homotopy-associative H-space, we can construct a 'projective 3-space over X', say B. The cohomology ring $H^*(B; Z_3)$ is then a truncated polynomial ring, in which the cube of the generator is non-zero. We now apply the Adem relations, as above, with $p = 3$.

It is natural to study H-spaces X such that $H^*(X; Z_p) = E(n; Z_p)$, because this is one of the simplest cohomology rings possible for an H-space. Our next theorem will provide examples of such H-spaces.

Let n be an odd integer greater than one; let us divide the primes p into two classes, P and Q. Let Z be the additive group of integers, and F the additive group of fractions t/u such that all the prime factors of u lie in Q.

It is easy to see that one can construct a countable CW-complex X

containing the n-sphere S^n as a subcomplex, so that X is a Moore space of type $Y(F, n)$, and so that the induced homomorphism

$$i_*: H_n(S^n) \to H_n(X)$$

realizes the embedding of Z in F. (Further details will be given below.) We have the theorem:

THEOREM. *Such a space X satisfies the following conditions.*

(i) $H^*(X; Z_p) = E(n; Z_p)$ *if* $p \in P$.

(ii) *If* $r \neq n$, *then* $\pi_r(X) \cong \sum_{p \in P} {}_p\pi_r(S^n)$, *where* ${}_p\pi_r(S^n)$ *denotes the p-component of the torsion group* $\pi_r(S^n)$.

(iii) *If* $2 \in Q$, *then X is a commutative H-space*; *if* (*in addition*) $3 \in Q$, *then X is homotopy-associative.*

We may summarize this theorem by the 'slogan' that 'an odd sphere is an H-space mod p'. The author originally made this remark in 1956, in order to answer a rather technical question concerned with the p-components of the homotopy groups of spheres. It was not published at that time. Taken in conjunction with Lemmas 1 and 2, it evidently yields examples of H-spaces of the following classes:

(i) commutative and not homotopy-associative;
(ii) commutative and homotopy-associative, but not equivalent-associative.

It also yields examples in connexion with the work of Stasheff and the author on A_n-spaces (**7**); however, no details of this application will be given in the present note.

Proof of the Theorem. We first make explicit the construction of X. Let d_m ($m = 1, 2, \ldots$) be a sequence of positive integers, to be chosen later; let e be a vertex in S^n and let $f_m: S^n, e \to S^n, e$ be a map of degree d_m. We may construct a CW-complex X by taking the product

$$\Big(\bigcup_m [2m-1, 2m]\Big) \times S^n$$

and identifying each point $(2m, x)$ with $\big(2m+1, f_m(x)\big)$. (X is thus a 'telescope'.) The sphere S^n is embedded in X as $1 \times S^n$. We easily see that

$$H_n(X) = F',$$

where F' is the additive group of fractions $t/(d_1 d_2 \ldots d_m)$; the induced homomorphism

$$i_*: H_n(S^n) \to H_n(X)$$

realizes the embedding of Z in F'. We can ensure that $F' = F$ by a

suitable choice of the integers d_m. By our construction, X is $(n-1)$-connected and has $H_r(X) = 0$ for $r \neq 0, n$; X is therefore a Moore space, of the sort required.

We establish part (ii) of the theorem by two applications of Serre's C-theory (**6**). Let C_Q be the class of torsion groups in which the order of each element is a product of primes in Q; similarly for C_P. Then the homomorphism

$$i_*\colon H_r(S^n) \to H_r(X)$$

is an isomorphism mod C_Q for each r. By the main theorems of (**6**) we see that

$$i_*\colon \pi_r(S^n) \to \pi_r(X)$$

is an isomorphism mod C_Q for all r.

Next, let R be the group of rationals; then the Eilenberg–Maclane groups $H_r(R, n)$ may be described as follows. $H^*(R, n;\, R)$ is an algebra over R on one generator of dimension n; it is a polynomial algebra or an exterior algebra according as n is even or odd. For $r > 0$, $H_r(R, n;\, Z)$ is the vector-space dual (over R) of $H^r(R, n;\, R)$. (These facts are easily established by induction over n.) Let Y be a space of type $K(R, n)$; then it is possible to find a map $f\colon X \to Y$ such that the homomorphism

$$f_*\colon H_n(X) \to H_n(Y)$$

realizes the embedding of F in R. Since n is odd,

$$f_*\colon H_r(X) \to H_r(Y)$$

is an isomorphism mod C_P for all r. Hence

$$f_*\colon \pi_r(X) \to \pi_r(Y)$$

is an isomorphism mod C_P for all r. These conclusions establish part (ii) of the theorem.

We begin work on part (iii) of the theorem by considering the symmetric square ΣX of X; a commutative product on X is equivalent to a retraction of ΣX onto X. We shall show that, if $2 \in Q$, then X is a deformation retract of ΣX, so that X admits one and only one homotopy class of commutative products.

For this purpose only the homotopy-type of the pair ΣX, X is relevant; and this is determined by the homotopy-type of X, which is determined by the data. We can therefore work in terms of any chosen X. Let us take X to be a 'simplicial telescope', in the obvious sense. Then X is an enumerable CW-complex, and X^2 admits a triangulation which is preserved when the two coordinates are permuted. In this way ΣX becomes a simplicial CW-complex. Moreover, X is the union of an increasing sequence of subcomplexes, each equivalent to

S^n; therefore ΣX is the union of an increasing sequence of subcomplexes, each equivalent to ΣS^n. Since n is odd, we have

$$H_r(\Sigma S^n) = \begin{cases} Z & (r = 0, n), \\ Z_2 & (r = n+2, n+4, \ldots, 2n-1), \\ 0 & \text{otherwise}. \end{cases}$$

In order to calculate $H_r(\Sigma X)$ as a direct limit we need only know the homomorphism

$$(\Sigma f_m)_*\colon H_r(\Sigma S^n) \to H_r(\Sigma S^n)$$

induced by a map $f_m\colon S^n \to S^n$ of degree d_m. This homomorphism is zero if $r > n$ and d_m is even, as one sees by the method of [(**8**) § 11]. We conclude that, if $2 \in Q$, then

$$H_r(\Sigma X, X) = 0$$

for each r. It follows that X is a deformation retract of ΣX.

We must now insert a slight digression. Since the homotopy-type of X is well-determined, we can prove part (i) of the theorem by calculating $H^*(X; Z_p)$ for some convenient X, for example, a telescope. More generally, let A be an abelian group and let G be a group in C_P; then we have

$$F \otimes_Z F = F, \qquad \operatorname{Tor}_Z(F, A) = 0,$$
$$\operatorname{Hom}_Z(F, G) = G, \qquad \operatorname{Ext}_Z(F, G) = 0.$$

For example, the group $\operatorname{Ext}_Z(F, G)$ can be found by calculating $H^{n+1}(X; G)$ when X is a telescope; or, equivalently, we can remark that the chain groups of such a telescope provide a resolution of F over Z.

We complete the proof of part (iii) of the theorem by comparing X with a loop-space. Let S^2X be the (reduced) double suspension of X, taking e as the base-point; S^2X is thus a Moore space of type $Y(F, n+2)$. We can embed $S(\Sigma X)$ in $\Sigma(SX)$ by the rule

$$i(t, (x, y)) = ((t, x), (t, y)) \quad (0 \leqslant t \leqslant 1).$$

This is consistent with the embedding of SX in each. Therefore we can embed $S^2(\Sigma X)$ in $\Sigma(S^2X)$ consistently with the embedding of S^2X in each; we see that

$$H_r(\Sigma S^2X, S^2\Sigma X) = 0.$$

The retraction from ΣX to X gives (by double suspension) a retraction from $S^2\Sigma X$ to S^2X; we can now extend this to a retraction on ΣS^2X.

Let us write

$$r\colon \Sigma X \to X, \qquad R\colon \Sigma S^2X \to S^2X$$

for the retractions we have constructed, and let Ω^2S^2X be the double loop-space on S^2X. The product R on S^2X induces a product μ_R on

$\Omega^2 S^2 X$. We have an embedding

$$j\colon X \to \Omega^2 S^2 X;$$

we now show that j is a homomorphism, with respect to the products r, μ_R. The verification is trivial. For

$$\begin{aligned}
(jr(x_1,x_2))(t,u) &= (t,u,r(x_1,x_2)) \quad \text{(by definition of } j)\\
(\mu_R(jx_1,jx_2))(t,u) &= R((jx_1)(t,u),\,(jx_2)(t,u)) \quad \text{(by definition of } \mu_R)\\
&= R((t,u,x_1),\,(t,u,x_2))\\
&= Ri(t,u,(x_1,x_2))\\
&\qquad \text{(where } i \text{ is the embedding of } S^2\Sigma X \text{ in } \Sigma S^2 X)\\
&= (t,u,r(x_1,x_2)) \quad \text{(by construction of } R).
\end{aligned}$$

We have thus shown that j is a homomorphism provided that the product used in $\Omega^2 S^2 X$ is μ_R. On the other hand, this product in $\Omega^2 S^2 X$ is homotopic to the ordinary product of loops, as is well known; and the latter is homotopy-associative. We conclude that there is a homotopy

$$h\colon I \times X^3 \to \Omega^2 S^2 X$$

in $\Omega^2 S^2 X$, between the maps

$$\mu(\mu \times 1),\ \mu(1 \times \mu)\colon X^3 \to X.$$

It remains only to compress this homotopy into X.

Since the groups $\pi_r(X)$, $\pi_{r+2}(S^2X)$ are in C_P for $r \neq n$, the groups $\pi_r(\Omega^2 S^2 X, X)$ are also in C_P. It is easy to calculate $H_r(\Omega^2 S^2 X)$, $H^r(\Omega^2 S^2 X;\,R)$ in a limited range of dimensions; since n is odd, we find that

$$H_r(\Omega^2 S^2 X, X) = \begin{cases} 0 & \text{if } 0 \leqslant r < p(n+1)-2,\\ Z_p & \text{if } r = p(n+1)-2,\end{cases}$$

where p is the smallest prime in P. Therefore $\pi_r(\Omega^2 S^2 X, X) = 0$ if $r < p(n+1)-2$. On the other hand, if G is a group in C_P, we have

$$H^r(I \times X^3, (0 \cup 1)\times X^3;\,G) = 0$$

if $r > 3n+1$. If $p > 3$, we have $3n+1 < p(n+1)-2$, and the theory of obstructions shows that the required compression is possible. This completes the proof.

To give our next example, we retain some of the notations used above, so that p is the smallest prime in the set P; we suppose that $p > 2$. Let Y be a space of type $Y(F,n)$, as above, and let W be a space obtained from Y by the method of 'killing homotopy groups', retaining its first two non-zero groups (which lie in dimensions n, $n+2p-3$) and killing the remainder. Then W is not equivalent to any H-space W'

which is both associative and commutative; for by a theorem of J. C. Moore, A. Dold, and R. Thom (**3**) such a space W' is weakly equivalent to a Cartesian product of Eilenberg–Maclane spaces, which W is not. (In fact, W has a non-zero k-invariant, given by the Steenrod operation P^1.)

However, W is weakly equivalent to a loop-space because the Postnikov invariant of W is stable; we therefore dispose of various devices for showing that W is equivalent to an associative H-space. We shall also show that W admits a commutative product. In what follows, we may suppose (as before) that our spaces are simplicial CW-complexes. We have a retraction from $W \cup \Sigma Y$ to W; we shall show that there is no obstruction to extending this retraction over ΣW. In fact, W may be obtained from Y by adjoining cells of dimension $n+2p+2$ and higher; therefore

$$H^r(\Sigma W, W \cup \Sigma Y; G) = 0$$

for $r < n+2p+2$. On the other hand, $\pi_r(W) = 0$ for $r > n+2p-3$. The extension is therefore possible.

Our further constructions depend on a modification of the 'method of killing homotopy groups'. As above, we suppose the primes p divided into two classes P and Q.

Lemma 3. *Suppose given an integer n and a CW-complex X such that $\pi_r(X)$ is a torsion group for $r \geqslant n$. Then there is a CW-complex Y containing X such that*

$$i_*\colon \pi_r(X) \to \pi_r(Y)$$

is an isomorphism mod C_Q *for all r, while*

$$\pi_r(Y) = 0 \bmod C_P \quad \textit{for } r \geqslant n.$$

Proof. Suppose that we form a space X' by attaching to X the cone on a $Y(G, n)$, where $G \in C_Q$. Then we have

$$\pi_r(X', X) = 0 \bmod C_Q$$

for all r. If we take
$$G = \sum_{p \in Q} {}_p\pi_n(X),$$
we can ensure that
$$\pi_n(X') = \sum_{p \in P} {}_p\pi_n(X).$$

This gives the first step of an obvious induction, proving the lemma.

It is easily shown (by obstruction-theory) that, if we start with a map $f\colon X \to X'$ between two spaces such as X, we can extend it to a map $g\colon Y \to Y'$.

Lemma 3 will provide us with all our remaining examples in a very simple fashion. Let us begin from the 7-sphere S^7; it admits various

products, but they are neither homotopy-associative nor homotopy-commutative. For each product μ, the separation element which shows that μ is not homotopy-associative lies in $\pi_{21}(S^7)$. Let Q be the set of primes p such that ${}_p\pi_{21}(S^7) \neq 0$; then Q is finite. Using Lemma 3, we may kill the corresponding p-components from dimension 21 upwards, obtaining a space X. Any product $\mu\colon S^7 \times S^7 \to S^7$ can be extended to a product $\mu'\colon X \times X \to X$; choose one such product μ'. By construction, we have an associating homotopy defined over $S^7 \times S^7 \times S^7$; and the further obstructions to extending this homotopy over $X \times X \times X$ lie in cohomology groups which vanish.

The H-space X is therefore homotopy-associative. It is not equivalent-associative because Lemma 1 applies for an infinity of primes p. It is not homotopy-commutative because we have done nothing to alter the separation elements in $\pi_{14}(S^7)$.

Similarly, let us begin with a product

$$\mu\colon S^7 \times S^7 \to S^7$$

such that the separation element to homotopy-commutativity is of order 2^k in $\pi_{14}(S^7)$. [It is easy to see that such a product exists; cf. (3.1) of (**5**)]. Using Lemma 3, we may kill the 2-components from dimension 14 upwards, obtaining a space X. Arguing as before, X is an H-space and is homotopy-commutative. It is not equivalent-commutative because ΣX cannot be retracted onto X, owing to the non-zero operation

$$Sq^2\colon H^7(\Sigma X; Z_2) \to H^9(\Sigma X; Z_2).$$

It is not homotopy-associative, by Lemma 2.

Similarly, we may employ both the above devices. Let X be a space obtained from S^7 by killing ${}_2\pi_r(S^7)$ for $r \geqslant 14$ and ${}_p\pi_r(S^7)$ for $r \geqslant 21$ and a suitable finite set of p; then X admits a product which is homotopy-commutative and homotopy-associative, but neither equivalent-commutative nor equivalent-associative.

Lastly, let X be a space obtained from S^7 by killing $\pi_r(S^7)$ for $r \geqslant 13$. Then the Postnikov system of X is stable, so X is weakly equivalent to an iterated loop-space $\Omega^m Y$ for as large an m as we please. In particular, X is equivalent-associative and homotopy-commutative. It is not equivalent-commutative, for the same reason as before:

$$Sq^2\colon H^7(\Sigma X; Z_2) \to H^9(\Sigma X; Z_2)$$

is non-zero.

If we add to these examples the standard H-space structures on S^1, S^3, and S^7, we obtain the ten examples required.

We have in fact done more than was needed. At various points we were committed to prove propositions of the following form: 'a certain H-space (X, e, μ) does not have the property P'. On each occasion, we have proved in addition that (X, e, μ') does not have the property P for any other product μ'.

REFERENCES

1. J. Adem, 'Relations on iterated reduced powers', *Proc. Nat. Acad. Sci. U.S.A.* 39 (1953) 636–8.
2. A. Dold and R. Lashof, 'Principal quasi-fibrations and fibre homotopy equivalence of bundles', *Illinois J. of Math.* 3 (1959) 285–305.
3. A. Dold and R. Thom, 'Quasifaserungen und unendliche symmetrische Produkte', *Annals of Math.* 67 (1958) 239–81.
4. I. M. James, *The ten types of H-spaces* (mimeographed).
5. —— 'Products on spheres', *Mathematika* 6 (1959) 1–13.
6. J.-P. Serre, 'Groupes d'homotopie et classes de groupes abéliens', *Annals of Math.* 58 (1953) 258–94.
7. J. Stasheff [to appear].
8. S. D. Liao, 'On the topology of cyclic products of spheres', *Trans. American Math. Soc.* 77 (1954) 520–51.

THE COLLEAGUE MATRIX, A CHEBYSHEV ANALOGUE OF THE COMPANION MATRIX

By I. J. GOOD (*Teddington*)

[Received 31 March 1960]

1. Summary

In the theory of canonical matrices the companion matrix plays a central role. It is pointed out here that a different matrix, resembling the companion matrix, can play much the same role. I call it the 'colleague matrix'. It has the same relation to trigonometric polynomials that the 'companion' has to ordinary polynomials. I do not know whether there is a wider class of analogous matrices of any interest to which the companion and colleague both belong.

2. Chebyshev polynomials

The following definitions and formulae are among those given, for example, in (4). They are repeated here for convenience of reference.

$$T_n(x) = \cos(n\cos^{-1}x) \quad (n = 0, 1, 2,...), \tag{1}$$

$$U_n(x) = (1-x^2)^{-\frac{1}{2}}\sin\{(n+1)\cos^{-1}x\} \quad (n = 0, 1, 2,...), \tag{2}$$

$$U_0(x) = 1, \quad U_1(x) = 2x, \quad U_2(x) = 4x^2-1,...,$$

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_2(x) = 2x^2-1,..., \tag{3}$$

$$C_n(x) = 2T_n(\tfrac{1}{2}x), \quad S_n(x) = U_n(\tfrac{1}{2}x), \tag{4}$$

$$T_{n+1}(x) = 2xT_n(x)-T_{n-1}(x), \tag{5}$$

$$U_{n+1}(x) = 2xU_n(x)-U_{n-1}(x), \tag{6}$$

$$T_n(x) = \tfrac{1}{2}U_n(x)-\tfrac{1}{2}U_{n-2}(x), \tag{7}$$

$$T_n(x) = xU_{n-1}(x)-U_{n-2}(x), \tag{8}$$

$$T'_n(x) = nU_{n-1}(x), \tag{9}$$

$$U_n(x) = \tfrac{1}{2}\{T_n(x)-T_{n+2}(x)\}/(1-x^2). \tag{10}$$

By a *T-series* we mean one of the form

$$a_0+a_1T_1(x)+a_2T_2(x)+...,$$

terminating or non-terminating; and we similarly define *U*-series, *C*-series, and *S*-series. When they terminate they may be called 'finite Chebyshev series', as in (3), and are expressible as trigonometric poly-

Quart. J. Math. Oxford (2), 12 (1961), 61–68.

nomials with argument $\cos^{-1}x$. They can also be expressed as ordinary polynomials, but for the purposes of numerical calculation it is better not to make the rearrangement in most situations. For a discussion of their importance for numerical work, see, for example, (**4**) and (**5**).

By using formulae (4), (7), and (10), we can effectively express T, U, C, and S series in terms of one another.

3. Relationship of Chebyshev polynomials to determinants and matrices

By expanding each of the determinants in terms of its first column and using the recurrence relations (5) and (6), we can prove at once that

$$T_n(x) = \begin{vmatrix} 2x & -1 & 0 & . & . & 0 & 0 \\ -1 & 2x & -1 & . & . & 0 & 0 \\ 0 & -1 & 2x & . & . & 0 & 0 \\ . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & 2x & -1 \\ 0 & 0 & 0 & . & . & -1 & x \end{vmatrix} \tag{11}$$

and

$$U_n(x) = \begin{vmatrix} 2x & -1 & 0 & . & . & 0 & 0 \\ -1 & 2x & -1 & . & . & 0 & 0 \\ . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & 2x & -1 \\ 0 & 0 & 0 & . & . & -1 & 2x \end{vmatrix}. \tag{12}$$

These two results are examples of the well-known fact that a triple-diagonal determinant can be computed by means of a sequence of trinomial recurrence relations. [Compare, for example, (**8**), 102; (**5**) 376; and (**7**) 35.] In formula (11) we may, of course, interchange the two corner elements x and $2x$ (but the x cannot be interchanged with any of the other $2x$'s).

Formula (12) can be re-expressed: $S_n(\lambda)$ is the characteristic polynomial $|\lambda\mathbf{I}-\mathbf{K}|$ of the matrix

$$\mathbf{K} = \begin{bmatrix} 0 & 1 & 0 & . & . & 0 & 0 \\ 1 & 0 & 1 & . & . & 0 & 0 \\ 0 & 1 & 0 & . & . & 0 & 0 \\ . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & 0 & 1 \\ 0 & 0 & 0 & . & . & 1 & 0 \end{bmatrix}.$$

(Up to this point I have said nothing new.)

It can now be easily proved by induction that

$$U_n(x)+a_1U_{n-1}(x)+...+a_n$$

$$=\begin{vmatrix} 2x & -1 & 0 & . & . & 0 & 0 \\ -1 & 2x & -1 & . & . & 0 & 0 \\ . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & 2x & -1 \\ a_n & a_{n-1} & a_{n-2} & . & . & -1+a_2 & 2x+a_1 \end{vmatrix}. \quad (13)$$

The inductive proof can be obtained by expanding the determinant in terms of its first column. [This result is closely related to (**3**).] We can replace $2x$ by x and the U's by S's. We can also prove that

$$T_n(x)+a_1T_{n-1}(x)+...+a_n$$

$$=\begin{vmatrix} x & -1 & 0 & . & . & 0 & 0 \\ -1 & 2x & -1 & . & . & 0 & 0 \\ 0 & -1 & 2x & . & . & 0 & 0 \\ . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & 2x & -1 \\ a_n & a_{n-1} & a_{n-2} & . & . & -1+a_2 & 2x+a_1 \end{vmatrix}. \quad (14)$$

This can be proved by expanding the determinant in terms of its first column, and making use of equations (13) and (8).

Formula (13) can be expressed in the form: the characteristic polynomial of the matrix

$$\mathbf{A}=\begin{bmatrix} 0 & 1 & 0 & . & . & 0 & 0 \\ 1 & 0 & 1 & . & . & 0 & 0 \\ . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & 0 & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & . & . & 1-a_2 & -a_1 \end{bmatrix} \quad (15)$$

is $$a(\lambda) = S_n(\lambda)+a_1S_{n-1}(\lambda)+...+a_n.$$

Now it is well known that the characteristic polynomial of the matrix

$$\mathbf{B}=\begin{bmatrix} 0 & 1 & 0 & . & . & 0 & 0 \\ 0 & 0 & 1 & . & . & 0 & 0 \\ . & . & . & . & . & . & . \\ -b_n & -b_{n-1} & -b_{n-2} & . & . & -b_2 & -b_1 \end{bmatrix} \quad (16)$$

is $$b(\lambda) = \lambda^n+b_1\lambda^{n-1}+...+b_n,$$

and **B** is usually called the 'companion matrix' of this polynomial [see, for example, (**6**) 20]. Accordingly I shall call **A** the 'colleague matrix'

of $a(\lambda)$. Another possible name would be 'Chebyshev matrix', but this might be confused with a matrix of the form $T_n(\mathbf{A})$ as used in (**12**).

When the roots $\mu_1, \mu_2, \ldots, \mu_n$ of $b(\lambda) = 0$, are distinct, the companion matrix can be diagonalized by means of the similarity transformation $\mathbf{Q}^{-1}\mathbf{BQ}$, where $\mathbf{Q}$ is the Vandermonde or alternant matrix

$$\mathbf{Q} = \begin{bmatrix} 1 & 1 & . & . & 1 \\ \mu_1 & \mu_2 & . & . & \mu_n \\ . & . & . & . & . \\ \mu_1^{n-1} & \mu_2^{n-1} & . & . & \mu_n^{n-1} \end{bmatrix}.$$

[See, for example, (**6**) 73.] Likewise, if the roots $\lambda_1, \lambda_2, \ldots, \lambda_n$ of $a(\lambda)$, are distinct, the colleague matrix can be diagonalized by means of the similarity transformation $\mathbf{P}^{-1}\mathbf{AP}$, where $\mathbf{P}$ is what may be called a 'Chebyshev alternant matrix',

$$\mathbf{P} = \begin{bmatrix} 1 & 1 & . & . & 1 \\ S_1(\lambda_1) & S_1(\lambda_2) & . & . & S_1(\lambda_n) \\ . & . & . & . & . \\ S_{n-1}(\lambda_1) & S_{n-1}(\lambda_2) & . & . & S_{n-1}(\lambda_n) \end{bmatrix}.$$

This can be proved by showing that the columns of $\mathbf{P}$ are eigenvectors of the colleague matrix. There is an extension to the case of equal roots, as for the companion matrix [(**8**) Chap. VI] by means of a 'Chebyshev confluent alternant matrix'. The analogue of the example given in [(**8**) 60] is as follows.

Consider a colleague matrix of order 6 having eigenvalues α, α, α, β, β, γ (i.e. one triple, one double, and one single). Then

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ -a_6 & -a_5 & -a_4 & -a_3 & 1-a_2 & -a_1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ S_1(\alpha) & S_1'(\alpha) & 0 & S_1(\beta) & S_1'(\beta) & S_1(\gamma) \\ S_2(\alpha) & S_2'(\alpha) & (1/2!)S_2''(\alpha) & S_2(\beta) & S_2'(\beta) & S_2(\gamma) \\ S_3(\alpha) & S_3'(\alpha) & (1/2!)S_3''(\alpha) & S_3(\beta) & S_3'(\beta) & S_3(\gamma) \\ S_4(\alpha) & S_4'(\alpha) & (1/2!)S_4''(\alpha) & S_4(\beta) & S_4'(\beta) & S_4(\gamma) \\ S_5(\alpha) & S_5'(\alpha) & (1/2!)S_5''(\alpha) & S_5(\beta) & S_5'(\beta) & S_5(\gamma) \end{bmatrix}$$

$$= \begin{bmatrix} \alpha & 1 & 0 & \beta & 1 & \gamma \\ \alpha S_1(\alpha) & \{\alpha S_1(\alpha)\}' & (1/2!)\{\alpha S_1(\alpha)\}'' & \beta S_1(\beta) & \{\beta S_1(\beta)\}' & \gamma S_1(\gamma) \\ \alpha S_2(\alpha) & \{\alpha S_2(\alpha)\}' & (1/2!)\{\alpha S_2(\alpha)\}'' & \beta S_2(\beta) & \{\beta S_2(\beta)\}' & \gamma S_2(\gamma) \\ . & . & . & . & . & . \\ \alpha S_5(\alpha) & \{\alpha S_5(\alpha)\}' & (1/2!)\{\alpha S_5(\alpha)\}'' & \beta S_5(\beta) & \{\beta S_5(\beta)\}' & \gamma S_5(\gamma) \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ S_1(\alpha) & S_1'(\alpha) & 0 & S_1(\beta) & S_1'(\beta) & S_1(\gamma) \\ S_2(\alpha) & S_2'(\alpha) & (1/2!)S_2''(\alpha) & S_2(\beta) & S_2'(\beta) & S_2(\gamma) \\ . & . & . & . & . & . \\ S_5(\alpha) & S_5'(\alpha) & (1/2!)S_5''(\alpha) & S_5(\beta) & S_5'(\beta) & S_5(\gamma) \end{bmatrix} \begin{bmatrix} \alpha & 1 & 0 & 0 & 0 & 0 \\ 0 & \alpha & 1 & 0 & 0 & 0 \\ 0 & 0 & \alpha & 1 & 0 & 0 \\ 0 & 0 & 0 & \beta & 1 & 0 \\ 0 & 0 & 0 & 0 & \beta & 1 \\ 0 & 0 & 0 & 0 & 0 & \gamma \end{bmatrix}.$$

4. Analogue of Danielewsky's method

One method for obtaining the characteristic polynomial of a matrix is to reduce it to companion form by means of a succession of simple similarity transformations, and one form of this method is due to Danielewsky. [See, for example, (**2**) 375.] Although Mr. J. H. Wilkinson tells me that Danielewsky's method, as usually described, is very liable to be unstable, it is of theoretical interest to note that it can be modified in order to obtain the colleague matrix, and thus to obtain the characteristic polynomial in the Chebyshev form.

As in (**2**), I shall exemplify the procedure by means of a 4×4 matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}.$$

Let

$$\mathbf{E}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ e_{21} & e_{22} & e_{23} & e_{24} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{E}_1^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

where $$e_{22} = 1/a_{12}, \quad e_{2j} = -a_{1j}/a_{12} \quad (j \neq 2).$$

Let $\mathbf{D}_1 = \mathbf{E}_1^{-1}\mathbf{A}\mathbf{E}_1$. The top row of $\mathbf{D}_1$ is $(0,1,0,0)$. Let its second row be $(d_{21}, d_{22}, d_{23}, d_{24})$. Let .

$$\mathbf{E}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ e_{31} & e_{32} & e_{33} & e_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{E}_2^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ d_{21}-1 & d_{22} & d_{23} & d_{24} \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

where

$$e_{33} = 1/d_{23}, \quad e_{31} = (1-d_{21})/d_{23}, \quad e_{3j} = -d_{2j}/d_{23} \quad (j \neq 1, 3).$$

Let $\mathbf{D}_2 = \mathbf{E}_2^{-1}\mathbf{D}_1\,\mathbf{E}_2$. The top row of $\mathbf{D}_2$ is (0, 1, 0, 0); the second row is (1, 0, 1, 0). Let the third row be $(d_{31}, d_{32}, d_{33}, d_{34})$. Let

$$\mathbf{E}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ e_{41} & e_{42} & e_{43} & e_{44} \end{bmatrix}, \qquad \mathbf{E}_3^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ d_{31} & d_{32}-1 & d_{33} & d_{34} \end{bmatrix},$$

where

$$e_{44} = 1/d_{34}, \qquad e_{42} = (1-d_{32})/d_{34}, \qquad e_{4j} = -d_{3j}/d_{34} \quad (j \neq 2, 4).$$

Then $\mathbf{D}_3 = \mathbf{E}_2^{-1}\mathbf{D}_2\,\mathbf{E}_3$ is of the required form.

The general formulae defining the e's are

$$e_{ij} = -d_{i-1,j}/d_{i-1,i} \quad (j \neq i-2, i),$$
$$e_{i,i-2} = (1-d_{i-1,i-2})/d_{i-1,i},$$
$$e_{i,1} = 1/d_{i-1,i}.$$

Small or vanishing denominators ('pivots') can be coped with as in (**2**).

5. Approximate zeros of a real function

Suppose that we wish to approximate to the zeros of a function in a finite interval, which may be taken as $(-1, 1)$ without real loss of generality. Let the function be approximated to (by known methods) by means of a terminating S-series. Then the real zeros of this approximating polynomial can be obtained as a subset of the eigenvalues of the colleague matrix. For methods of obtaining eigenvalues see, for example, (**1**), (**5**), (**9**).

This method of locating zeros is analogous to the well-known method of finding zeros of an ordinary polynomial by making use of the companion matrix. If the iteration for finding the eigenvalues uses column vectors, then the analogue is with Bernoulli's method; if it uses row vectors, the analogue is with Dietzold's method.

Trouble may arise as usual if two of the zeros of the original function are very close together. In order to detect possible trouble it would be of use to know something about the zeros of the derivative of the function. (Formula (9) can be used to get the derivative.) If, at a stationary point, the value of the function is small, then there is a danger that two zeros will be missed.

Since I have not yet tried these numerical methods, I am not yet claiming anything for them. The justification for mentioning them in the present paper is in order to indicate further analogies between the companion and colleague matrices.

6. Location of zeros of a polynomial

A method has been given by Parodi (**10**) for circumscribing the zeros of a polynomial

$$z^n + b_p z^{n-p} + \dots + b_{n-1} z + b_n,$$

and the proof of his result depends on the use of the companion matrix. His argument, which was a little more complicated than necessary, again has an analogue for finite Chebyshev series. In fact, *all the zeros of*

$$U_n(z) + a_p U_{n-p}(z) + a_{p+1} U_{n-p-1}(z) + \dots + a_n$$

lie in the closure of the union of the interiors of the curves $|z| < 1$ *and*

$$|U_{n-p+1}(z)+U_n(z)+a_p\,U_{n-p}(z)| < \{|a_n|+|a_{n-1}|+\ldots+|a_{p+2}|+|a_{p+1}-1|\}\,|U_{n-p}(z)|.$$

The main value of this result will presumably be for circumscribing the *real* zeros of polynomials since it is only on the real axis that Chebyshev series have numerical advantages over ordinary polynomials.

Proof. Take equation (13) with n replaced by $n-p+1$, and the bottom row replaced by

$$a_n,\ a_{n+1},\ \ldots,\ a_{p+2},\ -1+a_{p+1},\ 2x+\{a_p-2x+(U_n+U_{n-p-1})/U_{n-p}\}.$$

We find on using equation (6) that

$$\begin{vmatrix} 2x & -1 & 0 & . & . & 0 & 0 \\ -1 & 2x & -1 & . & . & 0 & 0 \\ . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & 2x & -1 \\ a_n & a_{n-1} & a_{n-2} & . & . & -1+a_{p+1} & a_p+(U_n+U_{n-p+1})/U_{n-p} \end{vmatrix}$$
$$= U_n+a_p\,U_{n-p}+a_{p+1}\,U_{n-p-1}+\ldots+a_n.$$

The result now follows by means of an application of a theorem for the non-vanishing of a determinant due, in its original form, to Lévy. [For a list of references see (**11**).] The theorem is that a matrix $\{m_{rs}\}$ is non-singular if it is dominated by its diagonal in the sense that

$$|m_{rr}| \geqslant \sum_{s}^{s\neq r} |m_{rs}| \quad (r = 1, 2, \ldots)$$

with strict inequality for at least one value of r.

REFERENCES

1. Anonymous, *Modern computing methods* (Her Majesty's Stationery Office, London, 1957).
2. R. A. Buckingham, *Numerical methods* (London, 1957).
3. C. W. Clenshaw, 'A note on the summation of Chebyshev series', *Math. Tables and other Aids to Computation*, 9 (1955) 118–20.
4. C. Lanczos, Introduction to *Tables of Chebyshev polynomials* $S_n(x)$ *and* $C_n(x)$ (National Bureau of Standards Appl. Math. Ser., 9, Washington, 1952).
5. —— *Applied analysis* (London, 1957).
6. C. C. MacDuffee, *The theory of matrices* (Berlin, 1933; reprint, New York, 1946).
7. L. Mirsky, *An introduction to linear algebra* (Oxford, 1955).

8. H. W. Turnbull and A. C. Aitken, *An introduction to the theory of canonical matrices* (London, 1932).

9. N. S. Mendelsohn, 'The computation of complex proper values and vectors of a real matrix with application to polynomials', *Math. Tables and other Aids to Computation,* 11 (1957) 91–94.

10. M. Parodi, 'Sur la localisation des zéros des polynomes lacunaires', *Bull. Sci. Math.* (2) 82 (1958) 67–72.

11. G. B. Price, Review of a paper by Massonet, *Math. Rev.* 8 (1947) 499.

12. C. Lanczos, 'Spectroscopic eigenvalue analysis', *J. Washington Acad. Sci.* 45 (1955) 315–23; and especially the review in *Math. Rev.* 17 (1956) 669.

THE SPIN REPRESENTATION OF THE COMPOUND AND INDUCED MATRICES OF AN ORTHOGONAL MATRIX

By A. O. MORRIS (*Aberystwyth*)

[Received 25 June 1960]

1. In this paper I show that the spin character of the compound and induced matrices of an orthogonal matrix can be written in terms of the characters of the original matrix. An explicit form is found for the spin character of the second compound and second induced matrix, but, for higher degrees, the results soon become elaborate, and cannot be explicitly formulated.

Before proceeding, I quote the results (**1**, **2**). If A is an orthogonal matrix of order n^2, where $n = 2\nu$ or $2\nu+1$, and the basic spin character of A is denoted by $\zeta(A)$, then

$$\zeta(A) = \prod_{r=1}^{\nu} (2\cos\tfrac{1}{2}\theta_r), \tag{1.1}$$

where $e^{\pm i\theta_r}$ $(r = 1, 2,..., \nu)$ are the variable characteristic roots of A. Further, if A and B are orthogonal matrices with m and n rows respectively, and

$$\phi(A) = \zeta(A) \quad \text{when } A \text{ is an even-rowed matrix,} \tag{1.2}$$

$$\phi(A) = \sqrt{2}\zeta(A) \quad \text{when } A \text{ is an odd-rowed matrix,} \tag{1.3}$$

then

$$\phi(A \dot{+} B) = \phi(A)\phi(B). \tag{1.4}$$

The cases of compound and induced matrices are dealt with separately.

2. Compound matrices

As usual, the rth compound matrix of a matrix A is denoted by $A^{(1^r)}$. We first prove the following result.

Theorem I. *If A is an n-rowed orthogonal matrix, then*

$$\phi(A^{(1^2)}) = 2^{\frac{1}{2}\nu}[\nu-1, \nu-2,..., 1] \quad \textit{when } n = 2\nu, \tag{2.1}$$

except when A has negative determinant, in which case

$$\phi(A^{(1^2)}) = 0, \tag{2.2}$$

and

$$\phi(A^{(1^2)}) = 2^{\frac{1}{2}(\nu+1)}[\nu-\tfrac{1}{2}, \nu-\tfrac{3}{2},..., \tfrac{1}{2}] \quad \textit{when } n = 2\nu+1. \tag{2.3}$$

Quart. J. Math. Oxford (2), 12 (1961), 69–77.

The cases in which A has an even or odd number of rows are dealt with separately.

(i) ($n = 2\nu$) For an orthogonal matrix A of positive determinant, we write

$$A = A_1 \dot{+} A_2 \dot{+} \ldots \dot{+} A_\nu, \tag{2.4}$$

where

$$A_r = \begin{bmatrix} e^{i\theta_r} & \\ & e^{-i\theta_r} \end{bmatrix}.$$

The second compound matrix of A is

$$A^{\{1^2\}} = \sum_{r=1}^{\nu} A_r^{\{1^2\}} \dot{+} \sum_{r,s=1}^{\nu} A_r \times A_s.$$

By (1.4), we have

$$\phi(A^{\{1^2\}}) = \phi\Big(\sum_{r=1}^{\nu} A_r^{\{1^2\}}\Big)\phi\Big\{\sum_{r,s=1}^{\nu}(A_r \times A_s)\Big\}$$

$$= \prod_{r=1}^{\nu} \phi(A_r^{\{1^2\}}) \prod_{r,s=1}^{\nu} \phi(A_r \times A_s).$$

But, by (1.3) and [(**1**) 329], we have

$$\phi(A_r^{\{1^2\}}) = \sqrt{2}, \qquad \phi(A_r \times A_s) = 2\cos\theta_r + 2\cos\theta_s.$$

Hence, it follows that

$$\phi(A^{\{1^2\}}) = 2^{\frac{1}{2}\nu} \prod_{r,s=1}^{\nu} (2\cos\theta_r + 2\cos\theta_s). \tag{2.5}$$

If $\Delta(x_i)$ denotes the product of differences of n variables $x_1, x_2, \ldots, x_n$, then

$$\Delta(x_i) = \prod_{i<j} (x_i - x_j) = |x_s^{n-t}|.$$

Multiply numerator and denominator by $\Delta(2\cos\theta_r)$, giving $\phi(A^{\{1^2\}})$ in the form

$$\phi(A^{\{1^2\}}) = 2^{\frac{1}{2}\nu} \frac{\Delta(4\cos^2\theta_r)}{\Delta(2\cos\theta_r)} = 2^{\frac{1}{2}\nu} \frac{|(4\cos^2\theta_t)^{\nu-s}|}{|(2\cos\theta_t)^{\nu-s}|}.$$

Using the relationship

$$2\cos n\theta = (2\cos\theta)^n + \sum_{r=1}^{[\frac{1}{2}n]} (-1)^r \frac{n(n-r-1)!}{(n-2r)!\, r!} (2\cos\theta)^{n-2r}$$

and adding and subtracting rows as necessary, we obtain

$$\phi(A^{\{1^2\}}) = 2^{\frac{1}{2}\nu} |C_t^{2(\nu-s)}| \,/\, |C_t^{(\nu-s)}|,$$

where $C_t^{(r)} = 2\cos r\theta_t$ $(r \neq 0)$, $C_t^{(0)} = 1$.

But, when $n = 2\nu$ [(**2**) 244],

$$[\lambda] = |C_t^{(\lambda_s+n-s)}| \,/\, |C_t^{(n-s)}|, \tag{2.6}$$

and thus we prove that

$$\phi(A^{\{1^2\}}) = 2^{\frac{1}{2}\nu}[\nu-1, \nu-2, \ldots, 1].$$

For an orthogonal matrix of negative determinant,

$$A_1 = \begin{bmatrix} 1 & \\ & -1 \end{bmatrix}, \quad \text{and} \quad \phi(A_1) = 0.$$

Hence $$\phi(A^{\{1^2\}}) = 0.$$

(ii) ($n = 2\nu+1$) An orthogonal matrix of positive determinant can be written in the form

$$A = \sum_{r=1}^{\nu} A_r \dot{+} [1]. \tag{2.7}$$

In this case, the second compound matrix of A is

$$A^{\{1^2\}} = \sum_{r=1}^{\nu} A_r^{\{1^2\}} \dot{+} \sum_{r,s=1}^{\nu} (A_r \times A_s) \dot{+} \sum_{r=1}^{\nu} A_r.$$

Hence, we have

$$\begin{aligned}\phi(A^{\{1^2\}}) &= \prod_{r=1}^{\nu} \phi(A_r^{\{1^2\}}) \prod_{r,s=1}^{\nu} \phi(A_r \times A_s) \prod_{r=1}^{\nu} \phi(A_r) \\ &= 2^{\frac{1}{2}(\nu+1)} \frac{|C_t^{2(\nu-s)}|}{|C_t^{(\nu-s)}|} \prod_{r=1}^{\nu} (2\cos\tfrac{1}{2}\theta_r).\end{aligned}$$

Multiplying the numerator and denominator by

$$\prod_{r=1}^{\nu} (2\sin\tfrac{1}{2}\theta_r)$$

and simplifying the resulting determinant by the usual method gives

$$\phi(A^{\{1^2\}}) = 2^{\frac{1}{2}(\nu+1)} |S_t^{(2\nu-2s+1)}| \,/\, |S_t^{(\nu-s+\frac{1}{2})}|,$$

where $S_t^{(r)} = 2\sin r\theta_t$. But, we have [(**2**) 245]

$$[\lambda] = |S_t^{(\lambda_s+\nu-s+\frac{1}{2})}| \,/\, |S_t^{(\nu-s+\frac{1}{2})}|, \tag{2.8}$$

and hence $$\phi(A^{\{1^2\}}) = 2^{\frac{1}{2}(\nu+1)}[\nu-\tfrac{1}{2}, \nu-\tfrac{3}{2}, \ldots, \tfrac{1}{2}].$$

An orthogonal matrix of negative determinant takes the form

$$A = \sum_{r=1}^{\nu} A_r \dot{+} [-1]. \tag{2.9}$$

In this case, we have

$$A^{\{1^2\}} = \sum_{r=1}^{\nu} A_r^{\{1^2\}} \dot{+} \sum_{r,s=1}^{\nu} (A_r \times A_s) \dot{+} \sum_{r=1}^{\nu} (-A_r),$$

and hence $$\phi(A^{\{1^2\}}) = 2^{\frac{1}{2}(\nu+1)} \frac{|C_t^{2(\nu-s)}|}{|C_t^{(\nu-s)}|} \prod_{r=1}^{\nu} (2\sin\tfrac{1}{2}\theta_r).$$

Multiplying the numerator and denominator by

$$\prod_{r=1}^{\nu} (2\cos\tfrac{1}{2}\theta_r)$$

and simplifying in the usual way gives

$$\phi(A^{\{1^2\}}) = 2^{\frac{1}{2}(\nu+1)}|S_l^{(2\nu-2s+1)}| \,/\, |C_l^{(\nu-s+\frac{1}{2})}|.$$

By using the corresponding formula to (2.8) for negative determinant, i.e.

$$[\lambda] = |S_l^{(\lambda_s+\nu-s+1)}| \,/\, |C_l^{(\nu-s+\frac{1}{2})}|, \tag{2.10}$$

we prove that $\qquad \phi(A^{\{1^2\}}) = 2^{\frac{1}{2}(\nu+1)}[\nu-\frac{1}{2}, \nu-\frac{3}{2},..., \frac{1}{2}].$

As mentioned previously, even for the case of the third compound matrix the above process becomes very elaborate, and nothing will be gained in continuing with this method.

The following lemma is now proved:

LEMMA. (i) *If* $A = \sum_{r=1}^{\nu}[A_r]$, *then*

$$(a) \quad A^{\{1^m\}} = \sum_{s=0}^{\mu}\left\{\binom{\nu-2s}{\mu-s}\sum A_{1,2,...,2s}\right\} \quad (m = 2\mu), \tag{2.11}$$

$$(b) \quad A^{\{1^m\}} = \sum_{s=0}^{\mu}\left\{\binom{\nu-2s-1}{\mu-s}\sum A_{1,2,...,2s+1}\right\} \quad (m = 2\mu+1) \tag{2.12}$$

where $A_{1,2,...,q}$ *denotes the direct product* $A_1 \times A_2 \times ... \times A_q$, *with*

$$A_r = \begin{bmatrix} e^{i\theta_r} & \\ & e^{-i\theta_r} \end{bmatrix}.$$

(ii) *If* $A = \sum_{r=1}^{\nu}[A_r]\dotplus[1]$, *then*

$$(a) \quad A^{\{1^m\}} = \sum_{s=0}^{\mu}\left\{\binom{\nu-2s}{\mu-s}\sum A_{1,2,...,2s}\right\}\dotplus\sum_{s=1}^{\mu}\left\{\binom{\nu-2s+1}{\mu-s}\sum A_{1,2,...,2s-1}\right\}$$
$$(m = 2\mu), \tag{2.13}$$

$$(b) \quad A^{\{1^m\}} = \sum_{s=0}^{\mu}\left\{\binom{\nu-2s-1}{\mu-s}\sum A_{1,2,...,2s+1}\dotplus\binom{\nu-2s}{\mu-s}\sum A_{1,2,...,2s}\right\}$$
$$(m = 2\mu+1). \tag{2.14}$$

For $A = \sum_{r=1}^{\nu}[A_r]\dotplus[-1]$, *then* $A^{\{1^m\}}$ *is given by* (2.13) *and* (2.14) *with* $\sum A_{1,2,...,2s-1}$ *and* $\sum A_{1,2,...,2s}$ *replaced by* $\sum -A_{1,2,...,2s-1}$ *and* $\sum -A_{1,2,...,2s}$ *respectively.*

If A and B are any matrices, then the mth compound matrix of their direct sum can be written (**2**)

$$[A\dotplus B]^{\{1^m\}} = [A]^{\{1^m\}}\dotplus[B]^{\{1^m\}}\dotplus\sum_{r=1}^{m-1}([A]^{\{1^r\}}\times[B]^{\{1^{m-r}\}}).$$

Similarly, the mth compound matrix of the direct sum of ν matrices $\sum_{r=1}^{\nu}[A_r]$ is written

$$\Big[\sum_{r=1}^{\nu}[A_r]\Big]^{\{1^m\}} = \sum_{r=1}^{\nu}[A_r]^{\{1^m\}} \dotplus \sum_{(\lambda)}\sum([A_1]^{\{1^{\lambda_1}\}}\times[A_2]^{\{1^{\lambda_2}\}}\times\ldots\times[A_s]^{\{1^{\lambda_s}\}}), \tag{2.15}$$

where $\lambda_1+\lambda_2+\ldots+\lambda_s = m$, and the first summation is taken over all the partitions of m into s or less parts, and the second over the $\nu!/(\nu-s)!$ permutations of $A_1, A_2,\ldots, A_\nu$ taken s at a time.

If $A_r = \begin{bmatrix} e^{i\theta_r} & \\ & e^{-i\theta_r}\end{bmatrix}$, then we have

$$[A_r]^{\{1^p\}} = \begin{cases} 1 & (p=2), \\ 0 & (p>2). \end{cases}$$

Thus, for the case (i), we need consider only partitions of the form $(1^{2s}\,2^{\mu-s})$ for (a), and $(1^{2s+1}\,2^{\mu-s})$ for (b); and, for the case (ii), partitions of the form $(1^{2s-1}\,2^{\mu-s})$ and $(1^{2s}\,2^{\mu-s})$ for (a), and $(1^{2s+1}\,2^{\mu-s})$ and $(1^{2s}\,2^{\mu-s})$ for (b). Hence, for example, in the case (i)(a), corresponding to the partition $(1^{2s}\,2^{\mu-s})$, we have the term

$$\sum\{A_1\times A_2\times\ldots\times A_{2s}\}\times A_{2s+1}^{\{1^2\}}\times\ldots\times A_{s+\mu}^{\{1^2\}} = \sum A_1\times A_2\times\ldots\times A_{2s}.$$

The same term is obtained for the

$$\binom{\nu-2s}{\mu-s}$$

possible ways of obtaining $A_{2s+1},\ldots, A_{s+\mu}$ from the remaining $\nu-2s$ matrices $A_{2s+1}, A_{2s+2},\ldots, A_\nu$.

Hence the lemma follows, where we denote $A_1\times A_2\times\ldots\times A_q$ by $A_{1,2,\ldots,q}$.

Using this lemma, we prove the theorem:

THEOREM II. *If A is an orthogonal matrix, then $\phi(A^{\{1^m\}})$ can be expressed linearly in terms of the characters of A, with positive integral coefficients.*

By (1.4), and the above lemma, we need only show that $\phi(\sum A_{1,2,\ldots,p})$ can be expressed linearly in terms of the characters of A, with positive integral coefficients, and the theorem will follow immediately. We have

$$\phi(A_{1,2,\ldots,p}) = f(\cos\theta_1, \cos\theta_2,\ldots, \cos\theta_p),$$

which is symmetric in $\cos\theta_1,\ldots, \cos\theta_p$, and thus

$$\phi(\sum A_{1,2,\ldots,p}) = \prod\{f(\cos\theta_1, \cos\theta_2,\ldots, \cos\theta_p)\},$$

where the summation and product are taken over the $\binom{\nu}{p}$ possible combinations of $\cos\theta_1, \cos\theta_2,\ldots, \cos\theta_p$. Thus, $\phi(\sum A_{1,2,\ldots,p})$ is expressible in terms of monomial symmetric functions of $\cos\theta_1, \cos\theta_2,\ldots, \cos\theta_p$, with integral coefficients, and hence can be expressed linearly in terms of the characters of A, with integral coefficients. The theorem follows.

3. Induced matrices

The same process is now repeated for the induced matrices of A. We first find the explicit form of the spin character of the nth induced matrix of a two-rowed orthogonal matrix, and the spin character of the second induced matrix of an n-rowed orthogonal matrix. Secondly, the corresponding theorem to Theorem II for induced matrices is proved.

As usual, denote the nth induced matrix of A by $A^{\{n\}}$. We first prove the theorem:

THEOREM III. *If A is a two-rowed orthogonal matrix, then*

$$\phi(A^{\{m\}}) = \sqrt{2}\prod_{i=1}^{\nu}[i] \quad (m = 2\nu), \tag{3.1}$$

$$\phi(A^{\{m\}}) = \prod_{i=0}^{\nu}\left[\frac{2i+1}{2}\right] \quad (m = 2\nu+1). \tag{3.2}$$

If $A = \mathrm{diag}(e^{i\theta_r}, e^{-i\theta_r})$, then we have

$$A^{\{m\}} = \begin{cases} \mathrm{diag}(e^{mi\theta}, e^{-mi\theta}, e^{(m-2)i\theta}, e^{-(m-2)i\theta},\ldots, e^{2i\theta}, e^{-2i\theta}, 1) \\ \hfill (m = 2\nu), \\ \mathrm{diag}(e^{mi\theta}, e^{-mi\theta}, e^{(m-2)i\theta}, e^{-(m-2)i\theta},\ldots, e^{i\theta}, e^{-i\theta}) \\ \hfill (m = 2\nu+1). \end{cases}$$

Hence,

$$\phi(A^{\{m\}}) = \begin{cases} 2^{\frac{1}{2}(m+1)}\cos\tfrac{1}{2}m\theta\cos\tfrac{1}{2}(m-2)\theta\ldots\cos\theta & (m = 2\nu), \\ 2^{\nu}\cos\tfrac{1}{2}m\theta\cos\tfrac{1}{2}(m-2)\theta\ldots\cos\tfrac{1}{2}\theta & (m = 2\nu+1), \end{cases}$$

which give the above results.

We then have the theorem:

THEOREM IV. *If A is an n-rowed orthogonal matrix, then*

$$\phi(A^{\{2\}}) = 2^{\frac{1}{2}\nu}[\nu, \nu-1,\ldots, 1] \quad (n = 2\nu), \tag{3.3}$$

except when A has negative determinant, in which case

$$\phi(A^{\{2\}}) = 0, \tag{3.4}$$

and

$$\phi(A^{\{2\}}) = 2^{\frac{1}{2}(\nu+1)}[\nu+\tfrac{1}{2}, \nu-\tfrac{1}{2},\ldots, \tfrac{3}{2}] \quad (n = 2\nu+1). \tag{3.5}$$

The proof follows closely the proof of Theorem I, and thus only an outline is given here.

(i) $(n = 2\nu)$ A has the form (2.4), and hence the second induced matrix is

$$A^{(2)} = \sum_{r=1}^{\nu} A_r^{(2)} \dot{+} \sum_{r,s=1}^{\nu} (A_r \times A_s).$$

Therefore, the spin character of $A^{(2)}$ is

$$\phi(A^{(2)}) = 2^{\frac{1}{2}\nu} \prod_{r=1}^{\nu} (2\cos\theta_r) \prod_{r,s=1}^{\nu} (2\cos\theta_r + 2\cos\theta_s),$$

which by the usual method can be shown to give

$$\phi(A^{(2)}) = 2^{\frac{1}{2}\nu} |C_t^{(2\nu-2s+1)}| \,/\, |C_t^{(\nu-s)}|.$$

Hence, by (2.6), we have

$$\phi(A^{(2)}) = 2^{\frac{1}{2}\nu}[\nu, \nu-1, \ldots, 1].$$

For an orthogonal matrix of negative determinant, as in Theorem I,

$$\phi(A^{(2)}) = 0.$$

(ii) $(n = 2\nu+1)$ In this case, if A has positive determinant, it takes the form (2.7), and hence the second induced matrix of A is

$$A^{(2)} = \sum_{r=1}^{\nu} A_r^{(2)} \dot{+} \sum_{r,s=1}^{\nu} (A_r \times A_s) \dot{+} \sum_{r=1}^{\nu} A_r \dot{+} [1].$$

Hence, we have

$$\phi(A^{(2)}) = 2^{\frac{1}{2}(\nu+1)} \frac{|C_t^{(2\nu-2s+1)}|}{|C_t^{(\nu-s)}|} \prod_{r=1}^{\nu} (2\cos\tfrac{1}{2}\theta_r).$$

Multiplying the numerator and denominator by $\prod_{r=1}^{\nu} (2\sin\frac{1}{2}\theta_r)$ and simplifying as usual gives

$$\phi(A^{(2)}) = 2^{\frac{1}{2}(\nu+1)} |S_t^{(2\nu-2s+2)}| \,/\, |S_t^{\nu-s+\frac{1}{2}}|.$$

By (2.8), we prove that

$$\phi(A^{(2)}) = 2^{\frac{1}{2}(\nu+1)}[\nu+\tfrac{1}{2}, \nu-\tfrac{1}{2}, \ldots, \tfrac{3}{2}].$$

If A is of negative determinant, then $A^{(2)}$ has the form

$$A^{(2)} = \sum_{r=1}^{\nu} A_r^{(2)} \dot{+} \sum_{r,s=1}^{\nu} (A_r \times A_s) \dot{+} \sum_{r=1}^{\nu} (-A_r) \dot{+} [1],$$

and $\phi(A^{(2)})$ is shown to have the form (3.5) using the formula (2.10).

Finally, we prove the theorem:

Theorem V. *If A is an orthogonal matrix, then $\phi(A^{\{m\}})$ is expressible linearly in terms of the characters of A, with positive integral coefficients.*

The corresponding equation to (2.15) for the mth induced matrix is

$$A^{\{m\}} = \sum [A_r]^{\{m\}} \dotplus \sum_{(\lambda)} \sum ([A_1]^{\{\lambda_1\}} \times [A_2]^{\{\lambda_2\}} \times \ldots \times [A_s]^{\{\lambda_s\}}),$$

where $\lambda_1+\lambda_2+\ldots+\lambda_s = m$ and the first summation is taken over all partitions (λ) of m into s or less parts, and the second summation over all the $\nu!/(\nu-s)!$ permutations of $A_1, A_2,\ldots, A_\nu$, taken s at a time.

If A is an $(n = 2\nu)$-rowed orthogonal matrix, then let

$$A_r = \mathrm{diag}(e^{i\theta_r}, e^{-i\theta_r}).$$

Therefore, we have

$$\phi([A_1]^{\{\lambda_1\}} \times [A_2]^{\{\lambda_2\}} \times \ldots \times [A_s]^{\{\lambda_s\}}) = g(\cos\theta_1, \cos\theta_2,\ldots, \cos\theta_s),$$

and hence

$$\phi\{\sum ([A_1]^{\{\lambda_1\}} \times [A_2]^{\{\lambda_2\}} \times \ldots \times [A_s]^{\{\lambda_s\}}) = \prod \{g(\cos\theta_1, \cos\theta_2,\ldots, \cos\theta_s)\},$$

where the sum and product are taken over all the possible combinations of $\cos\theta_1$, $\cos\theta_2,\ldots, \cos\theta_s$. Hence,

$$\phi\{\sum ([A_1]^{\{\lambda_1\}} \times \ldots \times [A_s]^{\{\lambda_s\}})\}$$

is expressible in terms of monomial symmetric functions of $\cos\theta_1,\ldots,$ $\cos\theta_s$, with positive integral coefficients, and can be expressed linearly in terms of the characters of A. As for Theorem II, it follows immediately that $\phi(A^{\{m\}})$ can be expressed linearly in terms of the characters of A, with positive integral coefficients.

If A is an $(n = 2\nu+1)$-rowed orthogonal matrix of positive determinant, it can be written in the form

$$A = A_1 \dotplus [1],$$

where A_1 is an $(n = 2\nu)$-rowed orthogonal matrix. Hence we have

$$A^{\{m\}} = [1] \dotplus \sum_{p=1}^{m} A_1^{\{p\}}.$$

From the above, it follows immediately that $\phi(A^{\{m\}})$ is expressible linearly in terms of the characters of A, with positive integral coefficients.

If A is an $(n = 2\nu+1)$-rowed orthogonal matrix of negative determinant, then

$$A^{\{m\}} = [(-1)^m] \dotplus \sum_{p=1}^{m} (-1)^{m-p} A_1^{\{p\}},$$

and hence $\phi(A^{\{m\}}) = 0$ if m is odd, and $\phi(A^{\{m\}})$ can be expressed linearly n terms of the characters of A with integral coefficients, if m is even. Thus the theorem is proved.

Theorems II and V are important since they verify that a spin representation of a matrix representation of an orthogonal group is expressible in terms of the true and the spin representations of the orthogonal group. Whereas the theorems deal only with compound and induced matrices, the extension to the general case is obtainable from these by means of direct products and direct differences. Thus, the extension of the result presents no difficulty.

I wish to express my thanks to Professor D. E. Littlewood for his guidance in preparing part of this paper, and for his suggestions for improving the presentation of the paper.

REFERENCES

1. A. O. Morris, *J. London Math. Soc.* **33** (1958) 326–33.
2. D. E. Littlewood, *The theory of group characters and matrix representations of groups* (Oxford, 1950).

ON AN INEQUALITY IN PARTIAL AVERAGES

By J. F. C. KINGMAN (*Cambridge*)

[Received August 1960]

1. Atkinson, Watterson, and Moran (**1**) have proved the following inequality. If

$$a_{ij} \geqslant 0, \qquad p_i \geqslant 0, \qquad q_j \geqslant 0, \qquad \sum_i p_i = \sum_j q_j = 1,$$

and if we define

$$a_{i.} = \sum_j a_{ij} q_j, \qquad a_{.j} = \sum_i a_{ij} p_i, \qquad a_{..} = \sum_{i,j} a_{ij} p_i q_j,$$

then

$$\sum_{i,j} a_{ij} a_{i.} a_{.j} p_i q_j \geqslant a_{..}^3. \tag{1}$$

In the present paper I give a more direct proof of this result and I also show that it is a special case of a much more general inequality, which includes also certain inequalities conjectured by Moran.

The general inequality proved below is a consequence of the elementary inequality, valid for $X \geqslant 0$, $Y \geqslant 0$, $k \geqslant 1$, $\sum X > 0$,

$$\frac{\sum XY^k}{\sum X} \geqslant \left(\frac{\sum XY}{\sum X}\right)^k, \tag{2}$$

which follows from the convexity of $f(x) = x^k$.

2. I first give an alternative proof of (1). Write

$$S = \sum_{i,j} a_{ij} a_{i.} a_{.j} p_i q_j = \sum_{i,j,k} a_{ij} a_{ik} a_{.j} p_i q_j q_k.$$

Interchanging j and k, and adding the two resulting expressions for S we have

$$\begin{aligned} S &= \sum_{i,j,k} a_{ij} a_{ik} \tfrac{1}{2}(a_{.j} + a_{.k}) p_i q_j q_k \\ &\geqslant \sum_{i,j,k} a_{ij} a_{ik} (a_{.j} a_{.k})^{\frac{1}{2}} p_i q_j q_k \\ &= \sum_i p_i \Big\{ \sum_j a_{ij} a_{.j}^{\frac{1}{2}} q_j \Big\}^2 \\ &\geqslant \Big\{ \sum_i p_i \sum_j a_{ij} a_{.j}^{\frac{1}{2}} q_j \Big\}^2, \quad \text{by (2) with } k = 2, \\ &= \Big(\sum_j q_j a_{.j}^{\frac{3}{2}} \Big)^2 \\ &\geqslant \Big\{ \Big(\sum_j q_j a_{.j} \Big)^{\frac{3}{2}} \Big\}^2, \qquad \text{by (2) with } k = \tfrac{3}{2}, \\ &= a_{..}^3. \end{aligned}$$

Equality is attained when the three inequalities in the proof become equalities. It is readily seen that this occurs if and only if, for each i, either $p_i = 0$ or $a_{i.} = a_{..}$, and, for each j, either $q_j = 0$ or $a_{.j} = a_{..}$.

3. We may regard a_{ij}, $a_{i.}$, $a_{.j}$, and $a_{..}$ as partial averages of the a_{ij} over different sets of indices. We may generalize this concept as follows. Let $a_{ijk...}$ be a collection of non-negative numbers, and denote the index set $(ijk\ldots)$ by α. If we have non-negative numbers $p_i, q_j, \ldots$ satisfying

$$\sum_i p_i = \sum_j q_j = \ldots = 1$$

and if (β, β') is a partition of α into two subsets, we may write

$$p_i q_j \ldots = P_\alpha = Q_\beta Q'_{\beta'}.$$

Then we define the *partial average* of a_α over β' by

$$A_\alpha(\beta) = \sum_{\beta'} a_\alpha Q'_{\beta'}. \tag{3}$$

For example, if $\alpha = (ij)$, $\beta = (i)$,

$$A_{ij}(\beta) = \sum_j a_{ij} q_j = a_{i.}.$$

Thus, if $\beta = \alpha$, $A_\alpha(\beta) = a_\alpha$, and, if β is empty,

$$A_\alpha(\beta) = \sum_\alpha a_\alpha P_\alpha = \bar{a} \quad \text{(say)}.$$

We note that, for any $r > 0$,

$$\begin{aligned}\sum_\alpha a_\alpha P_\alpha A^r_\alpha(\beta) &= \sum_\beta Q_\beta A^r_\alpha(\beta) \sum_{\beta'} a_\alpha Q'_{\beta'}, \quad \text{since } A_\alpha(\beta) \text{ depends only on those indices in } \beta,\\ &= \sum_\beta Q_\beta A^{r+1}_\alpha(\beta) \geqslant \Big\{\sum_\beta Q_\beta A_\alpha(\beta)\Big\}^{r+1}, \quad \text{from (2)},\\ &= \bar{a}^{r+1}.\end{aligned}$$

We are now in a position to prove the theorem:

THEOREM. *Let $A_\alpha(n)$ $(n = 1, 2, \ldots)$ be partial averages of a_α, and let λ_n be non-negative numbers. Then*

$$S = \sum_\alpha a_\alpha P_\alpha \prod_n A^{\lambda_n}_\alpha(n) \geqslant \bar{a}^{1+\Sigma\lambda_n}. \tag{4}$$

Proof. Let $\Lambda = \sum_n \lambda_n$ and let $\kappa_n = \lambda_n/\Lambda$. We need only consider the case $\bar{a} > 0$, $\Lambda > 0$ since otherwise the inequality is trivially true. Then, since

$$\prod_n A^{\kappa_n}_\alpha(n) = \lim_{r\to 0}\Big\{\sum_n \kappa_n A^r_\alpha(n)\Big\}^{1/r},$$

$$S = \lim_{r\to 0} S_r,$$

where

$$S_r = \sum_\alpha a_\alpha P_\alpha\Big(\sum_n \kappa_n A^r_\alpha(n)\Big)^{\Lambda/r}.$$

For $0 < r < \Lambda$ we can apply (2) to give

$$\begin{aligned}\frac{S_r}{\bar{a}} &\geqslant \Big\{\sum_{\alpha} a_\alpha P_\alpha \sum_n \kappa_n A_\alpha^r(n)/\bar{a}\Big\}^{\Lambda/r} \\ &= \Big\{\bar{a}^{-1}\sum_n \kappa_n\Big(\sum_\alpha a_\alpha P_\alpha A_\alpha^r(n)\Big)\Big\}^{\Lambda/r} \\ &\geqslant \Big\{\bar{a}^{-1}\sum_n \kappa_n \bar{a}^{1+r}\Big\}^{\Lambda/r} = \bar{a}^{\Lambda}.\end{aligned}$$

Letting $r \to 0$ we get $$S \geqslant \bar{a}^{1+\Lambda}.$$

Since this proof depends on a limiting process, the cases of equality cannot be identified directly.

By taking α to contain different numbers of indices, we obtain a series of inequalities, of which the first two are

$$\sum_{i,j} a_{ij}^{1+\lambda_1} a_{i\cdot}^{\lambda_2} a_{\cdot j}^{\lambda_3} p_i q_j \geqslant a_{\cdot\cdot}^{1+\Lambda}, \tag{5}$$

$$\sum_{i,j,k} a_{ijk}^{1+\lambda_1} a_{i\cdot\cdot}^{\lambda_2} a_{\cdot j\cdot}^{\lambda_3} a_{\cdot\cdot k}^{\lambda_4} a_{\cdot jk}^{\lambda_5} a_{i\cdot k}^{\lambda_6} a_{ij\cdot}^{\lambda_7} p_i q_j r_k \geqslant a_{\cdot\cdot\cdot}^{1+\Lambda}. \tag{6}$$

The inequalities conjectured by Moran are obtained by taking appropriate values for the λ_n. For example, with $\lambda_1 = \lambda_5 = \lambda_6 = \lambda_7 = 0$, $\lambda_2 = \lambda_3 = \lambda_4 = 1$ in (6),

$$\sum_{i,j,k} a_{ijk} a_{i\cdot\cdot} a_{\cdot j\cdot} a_{\cdot\cdot k} p_i q_j r_k \geqslant a_{\cdot\cdot\cdot}^4.$$

REFERENCE

1. F. V. Atkinson, G. A. Watterson, and P. A. P. Moran, 'A matrix inequality', *Quart. J. of Math.* (Oxford) (2) 11 (1960) 137–40.